National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

University of Alberta

NMOS Process Development and Characterization

by

Nelson Poon

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

Department of Electrical Engineering

Edmonton, Alberta

Fall 1995

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

THE AUTHOR HAS GRANTED AN IRREVOCABLE NON-EXCLUSIVE LICENCE ALLOWING THE NATIONAL LIBRARY OF CANADA TO REPRODUCE, LOAN, DISTRIBUTE OR SELL COPIES OF HIS/HER THESIS BY ANY MEANS AND IN ANY FORM OR FORMAT, MAKING THIS THESIS AVAILABLE TO INTERESTED PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE IRREVOCABLE ET NON EXCLUSIVE PERMETTANT A LA BIBLIOTHEQUE NATIONALE DU CANADA DE REPRODUIRE, PRETER, DISTRIBUER OU VENDRE DES COPIES DE SA THESE DE QUELQUE MANIERE ET SOUS QUELQUE FORME QUE CE SOIT POUR METTRE DES EXEMPLAIRES DE CETTE THESE A LA DISPOSITION DES PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP OF THE COPYRIGHT IN HIS/HER THESIS. NEITHER THE THESIS NOR SUBSTANTIAL EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT HIS/HER PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE DU DROIT D'AUTEUR QUI PROTEGE SA THESE. NI LA THESE NI DES EXTRAITS SUBSTANTIELS DE CELLE-CI NE DOIVENT ETRE IMPRIMES OU AUTREMENT REPRODUITS SANS SON AUTORISATION.

ISBN 0-612-06527-8

Canada

University of Alberta

Library Release Form

Name of Author: Nelson Poon

Title of Thesis: NMOS Process Development and Characterization

Degree: Master of Science

Year this Degree Granted: 1995

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Nelson Poon

8606 - 156 Street

Edmonton, Alberta

Canada T5R 1Y4

August 22, 1995

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled NMOS Process Development and Characterization submitted by Nelson Poon in partial fulfillment of the requirements for the degree of Master of Science.

Steven Dew

Dr. S. K. Dew - Supervisor

Keith A Stromsmoe

Dr. K. A. Stromsmoe

M A Freeman

Dr. M. Freeman

August 11/95

To my sisters and parents

Abstract

A low cost, four mask n-channel metal-oxide-semiconductor (NMOS) integrated circuit process has been developed specifically for the capabilities of the Alberta Microelectronic Centre. Key features of this process include solid source diffusion doping, self-aligned plasma-enhanced chemical vapour deposited (PECVD) polysilicon gate, thermal (dry) oxide gate dielectric, PECVD silicon nitride-based LOCOS isolation, and single level aluminum metallization. Except for mask fabrication, all aspects from layout design to packaging and testing can be done in-house and are suitable for integration with optical or micromachined devices.

Transistors produced using this process have suitable electrical characteristics for use in digital and analog circuits. Typical (unadjusted) threshold voltages are V_T = 0.37 V and channel transconductances are $k' = 24 - 40\ \mu A/V^2$. Non-aggressive channel lengths of 5 to 40 microns have been demonstrated, but shorter lengths are likely possible without the requirement of ion implantation.

Acknowledgments

This research would not have been possible without the tremendous support of my supervisor, Dr. S. K. Dew. I would like to thank him and many others at the Alberta Microelectronic Centre for their help and encouragement throughout the course of this work. In particular, my thanks go out to Graham McKinnon, Doug Bowles, Tran Tran, Niall Tait, and, of course, Glen Fitzpatrick for his advice and assistance in film depositions. I also wish to thank Dr. N. G. Tarr of Carleton University for his much appreciated advice.

Table of Contents

1. INTRODUCTION 1

1.1 MOSFET THEORY 4

1.1.1 Qualitative Theory Discussion 5

1.1.2 Quantitative Square-Law Theory 9

2. NMOS FABRICATION OVERVIEW 15

3. FOUR MASK NMOS FABRICATION PROCESS 21

3.1 FILM GROWTH AND DEPOSITION 21

3.1.1 Thermal Growth of Silicon Dioxide 21

3.1.2 Plasma Enhanced Chemical Vapour Deposition (PECVD) 24

3.1.3 Sputtering (DC Planar Magnetron) 25

3.2 DOPANT DIFFUSION 26

3.3 ETCHING AND WET CHEMICAL CLEANING 29

3.3.1 Reactive Ion Etching (RIE) 29

3.3.2 RCA Cleaning 30

3.4 NMOS PROCESS DETAILS 32

3.4.1 Mask Design 32

3.4.2 Starting Wafers 34

3.4.3 Pad Oxidation 35

3.4.4 Silicon Nitride Deposition 39

3.4.5 Field Region Pattern Etch 40

3.4.6 Field Oxidation 44

3.4.7 Nitride and Pad Oxide Removal 47

3.4.8 Gate Oxidation 48

3.4.9 Polysilicon Gate Deposition 49

3.4.10 Gate Pattern Etch 50

3.4.11 Source/Drain Gate Oxide Etchback 51

3.4.12 Source/Drain Diffusion 51

3.4.13 Dielectric Deposition 55

3.4.14 Contact Pattern Etch 56

3.4.15 Front Metallization 57

3.4.16 Metal Pattern Etch 58

3.4.17 Backside Oxide Removal 59

3.4.18 Metal Anneal 60

3.4.19 Back Metallization 62

4. CHARACTERIZATION 63

4.1 WAFER LEVEL MEASUREMENTS 63

4.1.1 Nanospec/AFT Film Thickness Measurement System 63

4.1.2 Tencor Instruments Alpha-Step 200 64

4.1.3 Four Point Probe 65

4.1.4 Van der Pauw Measurements 67

4.1.5 Diffusion Junction Depth 69

4.1.6 Contact Resistance 70

4.1.6.1 Contact Strings 70

4.1.6.2 Four Terminal Kelvin Test Structure 72

4.2 DEVICE MEASUREMENTS 73

4.2.1 Diffused and Polysilicon Resistors 73

4.2.2 NMOS Transistors 73

4.2.2.1 DC Characteristics (I_D-V_{DS} Curves) 73

4.2.2.2 Sub-threshold Leakage Current (I_{D0}) 74

4.2.2.3 Threshold Voltage (V_{T0}) 75

4.2.2.4 Transconductance Parameter (k') 75

4.2.2.5 Channel Length Modulation Parameter (λ) 76

5. TEST RESULTS AND DISCUSSION 77

5.1 SHEET RESISTANCE AND JUNCTION DEPTH 78

5.1.1 Four Point Probe 80

5.1.2 Van der Pauw Measurements 82

5.1.3 Diffusion Junction Depth 88

5.1.4 Comparison against Diffusion Theory 89

5.2 CONTACT RESISTANCE 90

5.2.1 Contact Strings 90

5.2.2 Four Terminal Kelvin Test Structures 95

5.3 RESISTORS 98

5.4 TRANSISTORS 102

5.4.1 I_D - V_{DS} Curves 102

5.4.2 Sub-threshold Leakage Current (I_{D0}) 106

5.4.3 Threshold Voltage (V_{T0}) 106

5.4.4 Transconductance Parameter (k') 110

5.4.5 Channel Length Modulation Parameter (λ) 111

5.4.6 Comparison of I_D - V_{DS} Curves with SPICE 111

6. CONCLUSIONS 117

REFERENCES 120

List of Tables

Table 2-1. Four mask NMOS fabrication. 15

Table 3-1. PECVD silicon nitride deposition parameters. 40

Table 3-2. PECVD polysilicon deposition parameters. 49

Table 3-3. PECVD silicon dioxide deposition parameters. 56

Table 3-4. DC planar magnetron aluminum sputter conditions. 58

Table 5-1. Initial estimates for diffusion times. 80

Table 5-2. Four point probe measurements on diffusion test wafers. 80

Table 5-3. Van der Pauw sheet resistance of diffused crosses. 87

Table 5-4. Van der Pauw sheet resistance of doped polysilicon crosses. 87

Table 5-5. Diffusion calculations for wafer #1. 89

Table 5-6. Diffusion calculations for wafer #2. 90

Table 5-7. Contact resistance of wafer #1 using contact strings. 91

Table 5-8. Contact resistance of wafer #2 using contact strings. 91

Table 5-9. Calculated contact resistance using Kelvin test structures. 95

Table 5-10. Resistor values for fully processed wafers #1 and #2. 101

Table 5-11. Extrapolated threshold voltage for transistors 150 μm wide. 109

Table 5-12. Experimental values of k'. 110

List of Figures

Figure 1-1. A typical NMOS transistor.....2

Figure 1-2. Terminology, current flow direction, and biasing conditions for NMOS transistors.....5

Figure 1-3. NMOS transistor operation for $V_{GS} > V_T$ and increasing V_{DS}.....8

Figure 1-4. I_D-V_{DS} characteristic curves for an NMOS transistor.....9

Figure 2-1. NMOS fabrication steps.....16

Figure 2-2. Comparison between metal gate and polysilicon self-aligned gate.....19

Figure 3-1. Location of charges in thermally oxidized silicon.....24

Figure 3-2. Energy levels created by doping intrinsic silicon.....27

Figure 3-3. Four mask design on one mask substrate.....33

Figure 3-4. NMOS mask layout.....34

Figure 3-5. Wafer placement during oxidation.....36

Figure 3-6. Bird's beak formation in field oxidation.....46

Figure 3-7. Wafer and doping source placement during predeposition.....53

Figure 4-1. Four point probe placement for sheet resistance measurements.....66

Figure 4-2. Van der Pauw test structures. (a) Diffused & polysilicon Greek crosses, (b) Diffused & polysilicon offset quadrate crosses. A=length of arm, S=width of central square, and D=width of arm.....68

Figure 4-3. Top view of diffused and polysilicon contact strings.....71

Figure 4-4. Cross sectional view of diffused and polysilicon contact strings.....71

Figure 4-5. Diffused and polysilicon four terminal Kelvin test structures. Metal connects pads 2 and 4. 72

Figure 4-6. Top view of designed NMOS transistors. 74

Figure 5-1. Numbering scheme for wafer die. 77

Figure 5-2. Diffused Greek crosses (a) wafer #1, (b) wafer #2. 83

Figure 5-3. Diffused offset quadrate crosses (a) wafer #1, (b) wafer #2. 84

Figure 5-4. Doped polysilicon Greek crosses (a) wafer #1, (b) wafer 2. 85

Figure 5-5. Doped polysilicon offset quadrate crosses (a) wafer #1, (b) wafer #2. 86

Figure 5-6. SEM photograph of junction depth on a diffused wafer. 88

Figure 5-7. Contact string measurements on wafer #1 (a) diffused string, (b) doped polysilicon string.....93

Figure 5-8. Contact string measurements on wafer #2 (a) diffused string, (b) doped polysilicon string.....94

Figure 5-9. Kelvin contact resistance on wafer #1 (a) diffused, (b) doped polysilicon...96

Figure 5-10. Kelvin contact resistance on wafer #2 (a) diffused, (b) doped polysilicon.....97

Figure 5-11. Wafer #1 resistors (a) diffused, (b) doped polysilicon. 99

Figure 5-12. Wafer #2 resistors (a) diffused, (b) doped polysilicon. 100

Figure 5-13. I-V characteristics for a transistor with L = 40 μm (wafer #1). 103

Figure 5-14. I-V characteristics for a transistor with L = 20 μm (wafer #1). 103

Figure 5-15. I-V characteristics for a transistor with L = 10 μm (wafer #1). 104

Figure 5-16. I-V characteristics for a transistor with L = 40 μm (wafer #2). 104

Figure 5-17. I-V characteristics for a transistor with L = 20 μm (wafer #2). 105

Figure 5-18. I-V characteristics for a transistor with L = 10 μm (wafer #2). 105

Figure 5-19. Saturation region transistor operation (W = 150 μm) on wafer #1. 107

Figure 5-20. Linear region transistor operation (W = 150 μm) on wafer #1. 107

Figure 5-21. Saturation region transistor operation (W = 150 μm) on wafer #2. 108

Figure 5-22. Linear region transistor operation (W = 150 μm) on wafer #2. 108

Figure 5-23. SPICE I_D - V_{DS} curves for a transistor with L = 40 μm (wafer #1)..... 113

Figure 5-24. SPICE I_D - V_{DS} curves for a transistor with L = 20 μm (wafer #1)..... 113

Figure 5-25. SPICE I_D - V_{DS} curves for a transistor with L = 10 μm (wafer #1)..... 114

Figure 5-26. SPICE I_D - V_{DS} curves for a transistor with L = 40 μm (wafer #2)..... 114

Figure 5-27. SPICE I_D - V_{DS} curves for a transistor with L = 20 μm (wafer #2)..... 115

Figure 5-28. SPICE I_D - V_{DS} curves for a transistor with L = 10 μm (wafer #2)..... 115

List of Symbols

MOSFET	Metal-oxide-semiconductor field effect transistor
W	Transistor channel width
L	Transistor channel length
NMOS	N-channel metal-oxide-semiconductor
I_D	Transistor drain current
V_{DS}, V_{GS}, V_S, V_{BS}	Transistor terminal voltages
V_T	Threshold voltage
V_{DSAT}	Drain pinch off voltage
ϕ_F	Equilibrium electrostatic potential
k	Boltzmann constant
T	Temperature
q	Magnitude of the electron charge
n_i	Intrinsic carrier concentration
p	Majority mobile hole carrier concentration
N_A, N_D	Doping concentration

Q_B, Q_{ox}, Q_f, Q_m, Q_{it}, Q_{ot}	Electrical charge
ε_{Si}, ε_{ox}, ε_o	Dielectric constant
ϕ_{MS}	Work function difference
V_{T0}	Threshold voltage for $V_{BS} = 0$
C_{ox}	Oxide capacitance per unit area
t_{ox}	Oxide thickness
γ	Body effect coefficient
$\bar{\mu}_n$	Average mobility
k'	Transconductance parameter
λ	Channel length modulation parameter
PECVD	Plasma enhanced chemical vapour deposition
LOCOS	Local oxidation of silicon
B	Parabolic rate constant
B/A	Linear rate constant
t	Time
τ	Time offset

RF	Radio frequency
C_{S1}, C_{S2}	Surface doping concentration
x	Diffusion depth
D	Dopant diffusivity
Q	Dopant dose
RIE	Reactive ion etching
DI	De-ionized
TCA	1,1,1 - Trichloroethane
HMDS	Hexamethyldisilazane
BOE	Buffered oxide etch
PSG	Phosphosilicate glass
s	Four point probe inter-spacing
t	Sample thickness
R_s	Sheet resistance
d	Wafer boundary distance
A	Arm length

D	Arm width
S	Central square width
SEM	Scanning electron microscope
CBKR	Cross - bridge Kelvin resistor
R_c	Contact resistance
ρ_c	Specific contact resistance
A_c	Contact area
I_{D0}	Sub-threshold leakage current
$\overline{\rho}_s$	Average resistivity
x_j	Junction depth
C_B	Background doping concentration
$\overline{\sigma}_s$	Average conductivity

1. INTRODUCTION

Most transistors in today's microelectronics operate upon the field effect principle. The field effect is seen when the conductivity of a semiconductor is modulated by an applied electric field normal to the surface of the semiconductor. This principle was first described by J. E. Lilienfeld in a series of patents granted to him in the early 1930's [1,2]. Lilienfeld described the effect of applying a voltage to an aluminum gate electrode on the surface of an adjoining piece of cuprous sulfide (Cu_2S) semiconductor. He stated that the semiconductor's electrical conductivity would be influenced by the electrostatic force applied across the aluminum. Technological immaturity and the lack of practical semiconductor materials stalled the introduction of an operational field effect device until 1953. That year, G. C. Dacey and I. M. Ross announced the fabrication of the junction field effect transistor (JFET) [3]. Seven years later, in 1960, D. Kahng and M. M. Atalla built the world's first operational silicon metal-oxide-semiconductor field effect transistor (MOSFET) [4]. The MOSFET is basically a capacitor and two adjoining back to back p-n junctions. The capacitor is composed of a metal layer on top of a semiconductor with a thin dielectric oxide layer in between the two. The conductivity of the semiconductor is changed by applying a voltage on the metal layer. The change in conductivity in the semiconductor then controls the current flowing between the two

heavily doped regions of the p-n junctions. Modern MOSFETs are now commonly fabricated with polysilicon, instead of metal, for the top electrode or gate, thermally grown silicon dioxide as the oxide layer, and single crystal <100> plane n or p-type silicon as the semiconductor. N-type silicon is used to produce p-channel (hole current carrying) devices and p-type silicon is used for n-channel devices (electron current carrying). A typical n-channel MOSFET (NMOS) is pictured in figure 1-1. The two doped regions are known as the source and drain and the distance between the two is the channel length (L). The channel width (W) is also shown in the figure.

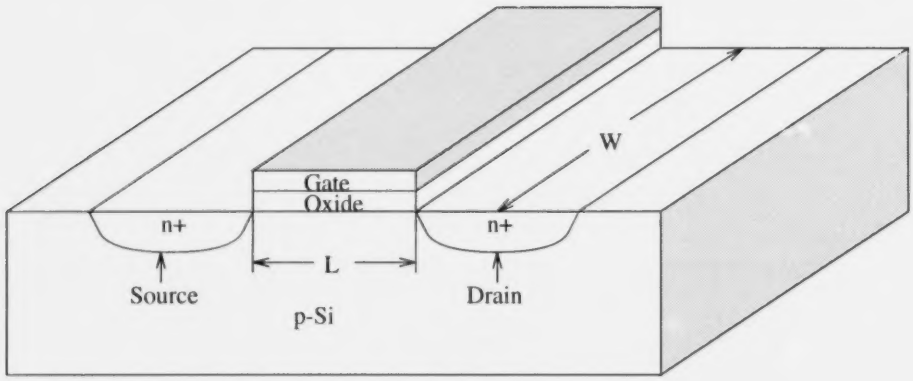


Figure 1-1. A typical NMOS transistor.

Since the 1960's, the MOSFET has become a key component in microelectronics technology. In 1963, F. M. Wanlass built an inverter by combining n and p-type MOSFETs [5]. This revolutionary new device consumed very low static

More recently, the MOSFET has led to the invention of devices such as dynamic random access memories (DRAM), erasable programmable read only memories (EPROM), and charge coupled devices (CCD). In addition, the MOSFET has been combined with the bipolar junction transistor (BJT) to form BIMOS (bipolar-n-channel MOSFET) and BICMOS (bipolar-CMOS) devices. This combination of the high drive capability of bipolar and the low power of CMOS can produce superior devices. The advanced integrated circuits of today routinely utilize CMOS or BICMOS technology, and MOSFETs with channel widths and lengths in the sub-micrometer range are now readily produced.

Because MOSFET technology encompasses such a wide range in microelectronics, the aim of this work will be to concentrate on one aspect of the MOSFET. This topic will be the fabrication and characterization of the n-channel MOSFET using the facilities available at the Alberta Microelectronic Centre (AMC). Equipment available at the AMC allows wet and dry oxidation, solid source diffusion doping, plasma enhanced chemical vapour deposition (PECVD), dry etching, and thin film sputtering. Thorough process and device characterization can also be done on-site. Ultimately, fabricated transistors could possibly be used in digital and analog circuits or integrated with micromachined devices under development at the AMC.

The remainder of this chapter will present a general discussion of MOSFET theory. Chapter 2 gives an overview of NMOS fabrication. A more detailed account

of the fabrication process is given in chapter 3. Characterization methods for fabricated transistors and test structures are given in chapter 4. Chapter 5 presents test results taken from fabricated devices. Chapter 6 concludes by providing brief closing remarks.

1.1 MOSFET THEORY

The material presented in this section provides a background to the dc (I_D-V_{DS}) characteristics of an n-channel MOSFET. The section first gives a qualitative discussion followed by a quantitative discussion based on the "square-law" theory. This theory gives two equations for the drain current, I_D. One is for transistor operation in the linear mode and the other for operation in the saturation mode. To simplify the discussion, the following figure shows the terminology, current flow direction, and usual biasing conditions normally used for NMOS transistors.

Normally, the source is grounded and the drain voltage is kept positive to keep the drain-substrate p-n junction reverse-biased. The voltage placed on the substrate (body) may be grounded or negative with respect to the source. This substrate voltage may be used to alter the device characteristics, as will be shown later. The NMOS transistor shown can be treated as a four terminal device.

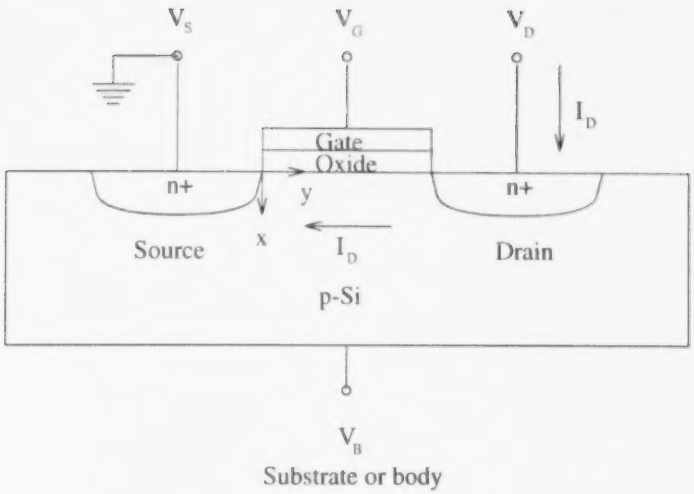


Figure 1-2. Terminology, current flow direction, and biasing conditions for NMOS transistors.

1.1.1 Qualitative Theory Discussion

To see how the voltage applied to the gate affects the drain-source current, assume that V_D is initially zero and the gate voltage is below a certain threshold voltage, V_T. This threshold voltage is the gate voltage required for the surface of the silicon to be inverted from p-type to n-type and marks the formation of a conducting n-channel (inversion layer) between the source and drain. Ideally, the gate threshold voltage required for NMOS is positive. However, because of non-ideal behaviour, a negative threshold voltage is sometimes encountered which implies that a conducting channel is present for $V_{GS} = 0$. Devices of this nature are called depletion mode

transistors. Transistors with no conducting channel at $V_{GS} = 0$ are called enhancement mode transistors. In either case, below the threshold voltage (sub-threshold), only a depletion layer exists at the surface, and no current can flow.

Increasing the gate voltage above the threshold voltage results in more and more electrons piling up at the silicon surface, and the conductance of the inversion layer rises accordingly. If it is now assumed that $V_{DS} > 0$, there exists a voltage drop between source and drain, and the conducting inversion layer between them acts as a simple resistor. A drain current proportional to V_{DS} begins to flow from the drain to source. This is known as the linear region of transistor operation.

With an increasing V_{DS}, the voltage drop between source and drain soon begins to reverse the effect of the positive gate voltage on the surface channel. Higher and higher drain voltages produce a wider and wider depletion region around the drain and the number of carriers nearest the drain drops. This reduces the conductance of the channel and the drain current levels off.

Increasing V_{DS} further eventually "pinches off" the inversion layer to a point. This pinch off voltage is denoted by V_{DSAT}. At drain voltages above pinch off, the pinched off part of the inversion layer enlarges from a point to a depleted region. The transistor now essentially acts as a constant current source and the drain current remains level. This region of operation is called the saturation region. Figure 1-3

graphically depicts the transistor operation described above for $V_{GS} > V_T$ and increasing V_{DS}.

The previous discussion assumes a single value for the gate voltage. To obtain a complete set of I_D - V_{DS} curves, the gate voltage is usually varied across a range. From the description given above, for an increasing gate voltage it can be expected that a corresponding increase in the channel conductance will occur. Graphically, this can be seen in the linear region as an increase in the slope of the I_D curve. In the saturation region, the increased channel conductance causes an increase in the drain voltage required for pinch off. A higher drain voltage is needed to reduce the increased number of electron carriers in the channel. The effects of V_{DS} and V_{GS} on I_D give rise to the dc characteristic curves shown in figure 1-4. Linear behaviour is seen left of the dashed line and saturation to the right.

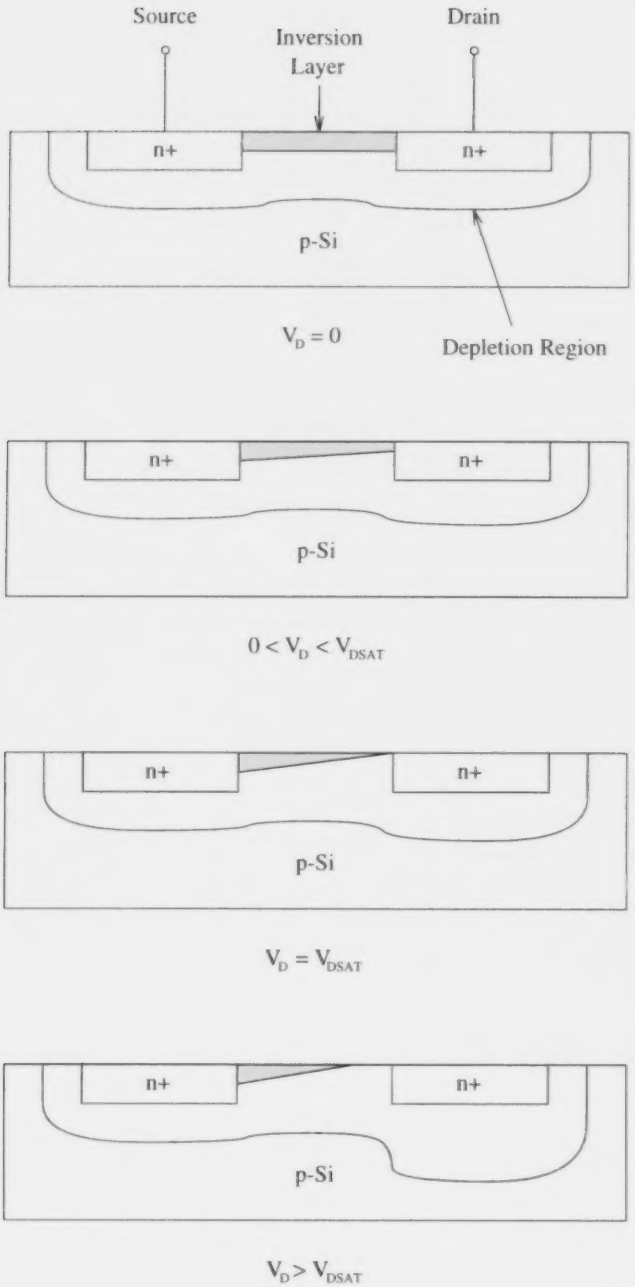


Figure 1-3. NMOS transistor operation for $V_{GS} > V_T$ and increasing V_{DS}.

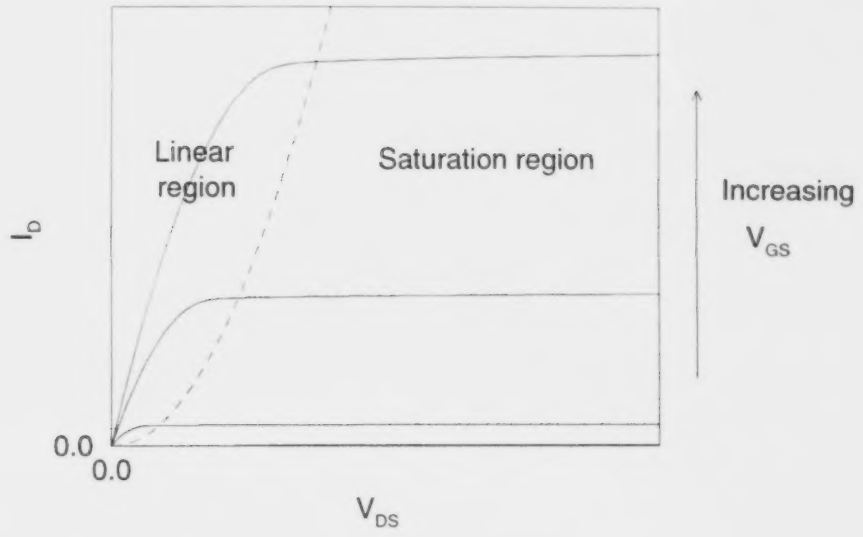


Figure 1-4. I_D-V_{DS} characteristic curves for an NMOS transistor.

1.1.2 Quantitative Square-Law Theory

Much of the literature on MOSFET theory includes a discussion on the square-law theory. It is often derived using semiconductor band theory or simple electrostatics principles. Both methods yield similar results and can be found in many references [6,7,8]. Therefore, a complete derivation will not be presented here. Instead the basic equations describing transistor operation with electrostatic principles will be given. These equations are similar to those derived by Hodges and Jackson [6].

Threshold Voltage

The threshold voltage is the most critical parameter to the transistor's electrical behaviour. The previous section defined the threshold voltage to be the gate voltage needed to form a surface inversion layer. This threshold voltage is composed of four separate terms. The first is the voltage required to change the semiconductor surface potential by $2|\phi_F|$. ϕ_F is the equilibrium electrostatic potential in a semiconductor. For a p-type semiconductor, ϕ_F is negative and is defined as

$$\phi_F = \frac{kT}{q} \times \ln \frac{n_i}{p} \qquad (1.1)$$

where k is Boltzmann's constant, T is temperature in degrees Kelvin, q is the magnitude of the electron charge, n_i is the intrinsic carrier concentration in the semiconductor, and p is the majority mobile carrier concentration (assumed to be equal to the semiconductor acceptor doping concentration, N_A). When the surface potential has changed by $2|\phi_F|$, the mobile electron density at the surface equals the original mobile hole density in the original substrate. With such an inversion layer present, a fixed charge is contained in the depletion region which is given by

$$Q_B = -\sqrt{2qN_A\varepsilon_{Si}\left|-2\phi_F - V_{BS}\right|} \qquad (1.2)$$

where ε_{Si} is the dielectric constant for silicon and V_{BS} is the substrate (or body) to source voltage. The second term of the threshold equation is the voltage required to offset this charge. With the addition of a negative substrate to source voltage it is

possible to increase the charge in the depletion region. This would, in turn, increase the threshold voltage which is useful since this allows direct control of the electrical properties.

The third term in the threshold voltage equation is dependent on the material used for the gate. It is denoted by ϕ_{MS} and is the work function difference between the gate material and bulk silicon. For NMOS with a doped polysilicon gate, $\phi_{MS} = \phi_{F(substrate)} - \phi_{F(gate)} = \frac{kT}{q}\ln\frac{n_i}{N_A} - \frac{kT}{q}\ln\frac{N_D}{n_i}$, where N_D is the donor doping concentration [6].

The last term defining the threshold voltage arises from the fact that there always exists an undesired charge, Q_{ox}, at the oxide-silicon interface. This charge arises from the formation of the gate oxide and will be discussed in more detail in chapter 3. Q_{ox} contributes a negative term to the threshold equation.

The threshold voltage equation can now be written as

$$V_{T0} = -2\phi_F + \phi_{MS} - \frac{Q_B}{C_{ox}} - \frac{Q_{ox}}{C_{ox}} \qquad (1.3)$$

V_{T0} is defined as the threshold voltage with $V_{BS} = 0$. C_{ox} is the gate oxide capacitance per unit area and is defined by $C_{ox} = \varepsilon_{ox}/t_{ox}$, where ε_{ox} is the dielectric constant of the oxide and t_{ox} is its thickness. For a non-zero V_{BS}, the following equation must be used

$$V_T = V_{T0} + \gamma(\sqrt{|-2\phi_F - V_{BS}|} - \sqrt{2|\phi_F|}) \qquad (1.4)$$

Gamma is known as the body effect coefficient and is defined as $\gamma = \sqrt{2q\varepsilon_{Si}N_A}/C_{ox}$.

I_D-V_{DS} Characteristics

Referring to figure 1.2, the voltage at a point y along the channel with respect to the source is V(y). If it is assumed that V_{GS} - V(y) is greater than V_T, the charge per unit area at y is given by

$$Q(y) = C_{ox}(V_{GS} - V(y) - V_T) \qquad (1.5)$$

Resistance dR of a length dy of the channel may be defined as

$$dR = \frac{dy}{W\overline{\mu}_n Q(y)} \qquad (1.6)$$

where $\overline{\mu}_n$ is the average (or effective) mobility of electrons in the channel. Because of scattering at the channel surface, the average mobility has a lower value than the bulk semiconductor mobility value which is, itself, a function of doping level. With dR defined, it can be said that the voltage drop across length dy is

$$dV = I_D dR = \frac{I_D}{W\overline{\mu}_n Q(y)} dy \qquad (1.7)$$

Substitution and algebraic manipulation gives

$$I_D dy = W\bar{\mu}_n C_{ox}(V_{GS} - V(y) - V_T)dV \tag{1.8}$$

Integrating the left side for y = 0 to L and the right side for V = 0 to V_{DS} gives

$$I_D = \frac{\bar{\mu}_n C_{ox} W}{2L}[2(V_{GS} - V_T)V_{DS} - V_{DS}^2] \tag{1.9}$$

which can be written

$$I_D = \frac{k'W}{2L}[2(V_{GS} - V_T)V_{DS} - V_{DS}^2] \tag{1.10}$$

where $k' = \bar{\mu}_n C_{ox}$ is known as the process transconductance parameter. The above equation is valid for $V_{GS} \geq V_T$ and $V_{DS} \leq V_{GS} - V_T$.

For the saturation region of operation, the above equation cannot be used. This is because at V_{DSAT}, the channel is pinched off and the charge at length L is zero. This implies $Q(L) = C_{ox}(V_{GS} - V_{DSAT} - V_T) = 0$. Therefore, $V_{DSAT} = V_{GS} - V_T$ and this gives the equation for the dashed transition line in figure 1-4. The equation for operation in the saturation region is obtained by substituting $V_{GS} - V_T$ for V_{DS}. Doing so results in

$$I_D = \frac{k'W}{2L}(V_{GS} - V_T)^2 \tag{1.11}$$

valid for $V_{GS} \geq V_T$ and $V_{DS} \geq V_{GS} - V_T$. This equation implies that the drain current is independent of V_{DS} in the saturation region. In reality, increasing V_{DS} above saturation slightly increases the drain current. This is due to electrostatic

coupling between the drain and mobile channel charge and shortening of the channel length with high V_{DS} as the depletion region expands horizontally. These effects on I_D can be approximated by an additional factor according to

$$I_D = \frac{k'W}{2L}(V_{GS} - V_T)^2(1 + \lambda V_{DS}) \qquad (1.12)$$

λ is known as the channel length modulation parameter.

2. NMOS Fabrication Overview

The fabrication process to produce the NMOS structure in figure 1-1 starts with a p-doped silicon wafer and essentially involves a repeated series of film deposition or film growth, patterning, and then etching or doping. Table 2-1 summarizes the four masks required for fabricating NMOS transistors. Figure 2-1 shows the result of each etch step.

Layer	Mask	Patterned Material
1	Active Device Region	Silicon nitride on pad oxide
2	Gate	Polysilicon on gate oxide
3	Contact	Dielectric oxide
4	Metal	Aluminum

Table 2-1. Four mask NMOS fabrication.

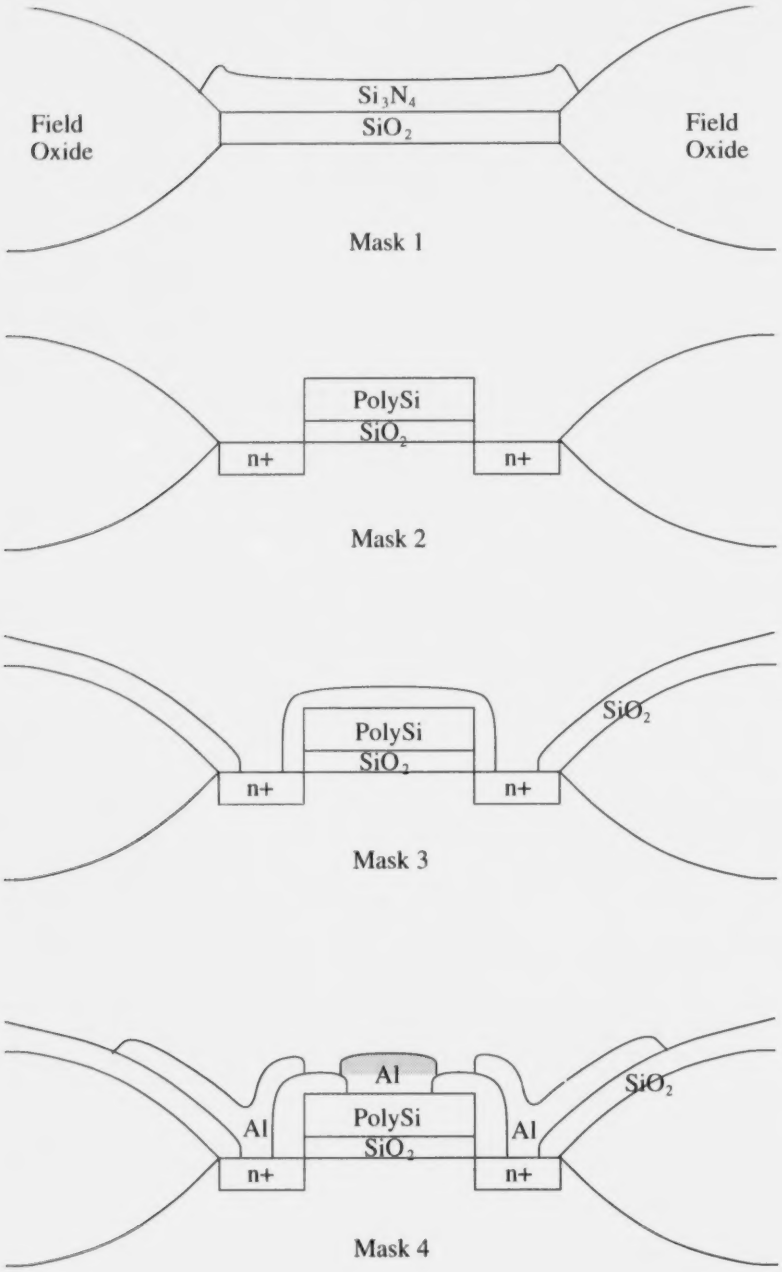


Figure 2-1. NMOS fabrication steps.

Film deposition in transistor fabrication is accomplished in many ways. The most common methods use physical vapour deposition (PVD), such as sputtering or evaporation, or chemical vapour deposition (CVD), such as plasma enhanced CVD (PECVD) or low pressure CVD (LPCVD).

The patterning of thin films is done by a step known as photolithography and involves the use of ultraviolet (UV) light to expose photoresist on a wafer. To define features on a wafer, a patterned chromium mask is placed over the wafer to be exposed. Chromium is opaque to UV light, so the pattern on the mask selectively exposes the resist. Upon exposure, the molecular structure of the photoresist changes. If the photoresist is a positive resist, exposure makes it more soluble in alkaline developer. Conversely, negative resists become insoluble in developer after exposure. After development, the formed pattern of photoresist acts as a mask for etching the underlying material. Etching is done with either wet chemical or dry etching. After etching, the photoresist is removed and the process may be repeated again for the next layer to be deposited.

One step in NMOS fabrication that is unique from the rest is semiconductor doping. This step forms the p-n^+ junctions required for NMOS devices. This may be done with ion implantation or diffusion. Dopant diffusion requires the wafer to be exposed to a liquid, gas, or solid source of dopant material at high temperature. This

is a low cost, high volume batch process. Ion implantation, however, is a capital intensive, highly controllable single wafer process.

The planar NMOS fabrication process, briefly described here, involves a minimum of four masking steps. The patterned layers build upon the previous layers to form the desired device. The starting point is to define the active device regions on the wafer. A bilayer composed of silicon nitride on silicon dioxide is first patterned on the wafer. The exposed silicon areas define the isolation regions. This use of nitride and oxide is part of the local oxidation of silicon (LOCOS) technique invented by E. Kooi in the early 1970's [9]. The LOCOS method isolates active regions with a thick, thermally grown field oxide without dramatically increasing surface non-planarity. Silicon nitride is resistant to oxidation and masks the device regions during this oxide growth. However, silicon nitride directly deposited on silicon suffers from significant stress which causes dislocations. The purpose of the thin underlying silicon dioxide layer, known as the pad oxide, is to relieve this stress. It also serves a dual purpose in providing a convenient etch stop for subsequent nitride removal. Ultimately, the pad oxide has too many trapped charges to be used as the gate oxide, so it must also be removed.

After the nitride and pad oxide are removed, the high quality gate oxide is grown and gate material is deposited. Then, the transistor gate region can be patterned. Instead of a metal, such as aluminum, it is desirable to use doped

polysilicon as the gate material. With polysilicon, a self-aligned gate process is used to form the gate and aid in source-drain doping [10]. In this process, the gate is patterned first and acts as a mask for the subsequent doping of the source and drain. Unlike aluminum, polysilicon can withstand the high temperatures associated with doping. The usual aluminum gate process defines the source and drain regions first. The gate must then overlap the source and drain to allow for mask registration errors and variation in dopant lateral diffusion. As well, a gate that does not extend completely across the length of the channel would make a device inoperable because an incomplete inversion layer would exist. The required overlap for the standard aluminum gate process gives rise to unwanted parasitic capacitances between the gate and drain and gate and source. No such overlap is needed in the self-aligned process. Figure 2-2 gives a comparison between the two methods.

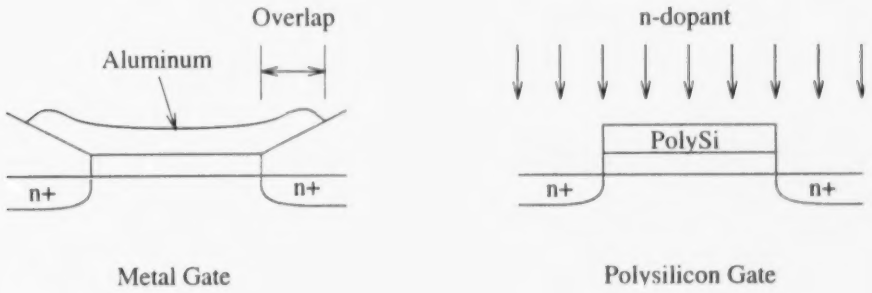


Figure 2-2. Comparison between metal gate and polysilicon self-aligned gate.

After formation of the gate and doping of the source and drain, a dielectric layer is deposited on the entire wafer. This provides electrical insulation between metal interconnections and silicon devices. The dielectric is most often CVD deposited silicon dioxide. Contact holes are then patterned, and the oxide is etched down to the polysilicon for gate contacts and down to the diffused regions for source and drain contacts.

The remaining step is to deposit, pattern, and etch the metallization layer. The metal layer makes the electrical contact through the contact holes, defines the required device interconnections, and provides the necessary metal pads for bonding or testing. A majority of the time, aluminum and aluminum alloys are used for metallization. Aluminum has good conductivity, is stable and adheres well to SiO_2, is easy to deposit using evaporation or sputtering, and has the ability to reduce native SiO_2 to ensure good contact to the silicon.

3. Four Mask NMOS Fabrication Process

The fabrication process briefly described in chapter 2 will now be discussed in more detail. To begin this chapter, background information on the individual processes of oxide growth, film deposition, dopant diffusion, etching, and wafer cleaning will be given. This will be followed by a precise description of the methods and techniques used for each fabrication step.

3.1 Film Growth and Deposition

3.1.1 Thermal Growth of Silicon Dioxide

When silicon is exposed to an oxygen ambient, a silicon dioxide layer forms on the surface. Silicon oxidation proceeds by the diffusion of an oxidant through the growing oxide layer. This oxidant then reacts with the underlying silicon at the $Si\text{-}SiO_2$ interface. Because of differences in molecular size, an oxide grown to a thickness t_{ox} consumes $0.45t_{ox}$ of the silicon underneath it and rises $0.55t_{ox}$ above the original silicon surface. Under room temperature, a "native" oxide grows to a self

limiting thickness of about 40 angstroms [11]. Under much higher temperatures, the oxide growth is greatly accelerated and becomes more durable and dense.

For the NMOS process, films of thermally grown silicon dioxide are required for the pad oxide, field oxide, and gate oxide. These films are grown in either an oxygen or steam ambient. With pure oxygen, the oxidation of silicon proceeds as

$$Si_{(s)} + O_{2(g)} \rightarrow SiO_{2(s)} \qquad (3.1)$$

In steam, the reaction proceeds as

$$Si_{(s)} + 2H_2O_{(g)} \rightarrow SiO_{2(s)} + 2H_{2(g)} \qquad (3.2)$$

The H_2O may be supplied by bubbling a carrier gas through water held near the boiling point, using an external steam generator, or by reacting hydrogen and oxygen gas in the oxidation tube (pyrogenic system).

Thermal oxidation rates are predicted using a model proposed by B. E. Deal and A. S. Grove [12]. The Deal - Grove model allows multiple oxidations to be considered and is quite accurate for oxides between 300 angstroms and 2 micrometers thick. For an oxide of thickness t_{ox}, the model is given by

$$t_{ox}^{\ 2} + At_{ox} = B(t + \tau) \qquad (3.3)$$

B is known as the parabolic rate constant and the ratio B/A is referred to as the linear rate constant. These constants are dependent on conditions such as temperature,

silicon wafer orientation, and wafer doping. The time required to grow to a thickness t_{ox} is t, and τ indicates the time offset required to correct for the presence of the initial oxide layer.

An unfortunate end result of silicon oxidation is the production of charges at or near the Si - SiO_2 interface. Four distinct charge sources are formed during silicon oxidation. These are fixed oxide charge (Q_f), mobile ionic charge (Q_m), interface trapped charge (Q_{it}), and oxide trapped charge (Q_{ot}) [13]. These four charges contribute to the charge Q_{ox} discussed in chapter 1. Fixed oxide charges are positive charges primarily due to structural defects (ionized silicon) in the oxide layer. Mobile ionic charges are due to ionic impurities such as Li^+ and Na^+. Interface trapped charges are negative or positive and originate from structural, oxidation induced defects, metal impurities, and other defects caused by bond breaking processes. Oxide trapped charges are holes or electrons which are trapped in the bulk of the oxide. Possible causes are from ionizing radiation (common in VLSI process equipment), avalanche injection, or some other similar processes. Fortunately, oxide trapped charges are usually annealed out at a low temperature. The location of these charges is shown in figure 3-1.

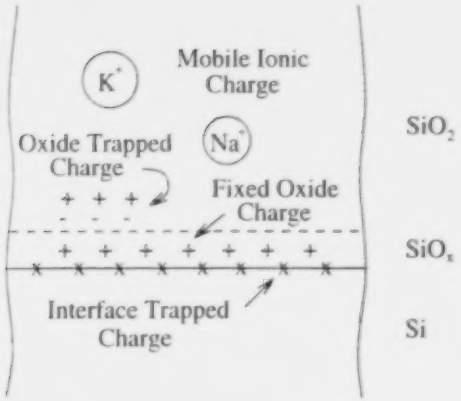


Figure 3-1. Location of charges in thermally oxidized silicon.

3.1.2 Plasma Enhanced Chemical Vapour Deposition (PECVD)

Chemical vapour deposition involves the use of either a gaseous compound that will decompose (pyrolyze) under heat to give a desired solid film or a mix of gases that will react to give the film. Substrate temperatures must usually be in the hundreds of degrees to allow the reaction to proceed. With the use of a plasma, the temperature of the substrate can be lowered significantly and thermal damage to the film reduced [14].

A plasma is formed when a low pressure gas is excited by a sufficiently high electric field. In PECVD, an RF voltage is used to form the plasma. Electrons in an RF field gain much more energy than ions or neutral molecules and, thus, have a

much higher effective temperature. These energetic electrons give molecules enough energy to react through inelastic collisions. These collisions produce free radicals which are neutral species whose electrons have gained extra energy. The free radicals are chemically very reactive and diffuse to the substrate where they react to form the desired film. Films commonly deposited by PECVD are silicon nitride and silicon dioxide. The non-equilibrium conditions cause films to normally be nonstoichiometric and often contain a small percentage of hydrogen.

3.1.3 Sputtering (DC Planar Magnetron)

Sputtering is a method of depositing films by using energetic ions to dislodge atoms from a solid target material. These mainly neutral atoms then condense onto the substrate to form a thin film. The target material is negatively biased using a DC or RF power supply and positive ions generated from a glow discharge sputter the target. Low pressure argon gas is commonly used as the source of ions and provides the medium for a glow discharge to exist. Argon is an inert gas and has a relatively good sputter yield. As the target material is bombarded with argon ions, secondary electrons are emitted. These secondary electrons then ionize neutral argon atoms and the glow discharge becomes self sustaining. The number of ions produced by secondary electron collisions is one of the limiting factors in the deposition rate.

To improve ion generation efficiency, magnets may be used to confine the electrons to a spiral path. This results in a longer electron path length within the high electric fields which improves efficiency. In planar magnetron sputtering, a series of magnets are strategically placed behind a planar target material such that the magnetic field lines leave the cathode, become nearly parallel to the target material, and then re-enter the cathode. In the region where the magnetic field is parallel to the cathode it is also perpendicular to the existing electric field. This confines the ionization region of the electrons close to the cathode and improves sputtering efficiency. The resulting area of target erosion is known as the racetrack due to the placement of the magnets to improve target utilization.

3.2 Dopant Diffusion

The source and drain of the MOSFET is heavily doped by using an appropriate donor or acceptor atom. For silicon, donor atoms from group VA of the periodic table are used. Acceptor atoms come from group IIIA. Doping of silicon with these atoms introduces additional energy levels within the intrinsic silicon band gap. Donor (n-type) atoms produce an additional energy level, E_d, close to the conduction band. Acceptor (p-type) atoms create an acceptor level, E_a, close to the valence band. The number of charge carriers in the silicon increases because the

energy barrier to form a charge carrier is reduced from the intrinsic band gap value (E_g) of 1.1 eV. Relative to the valence band, for n-type doping the energy barrier becomes E_g - E_d. For p-type doping, it becomes E_a. Figure 3-2 shows the energy levels introduced by doping intrinsic silicon.

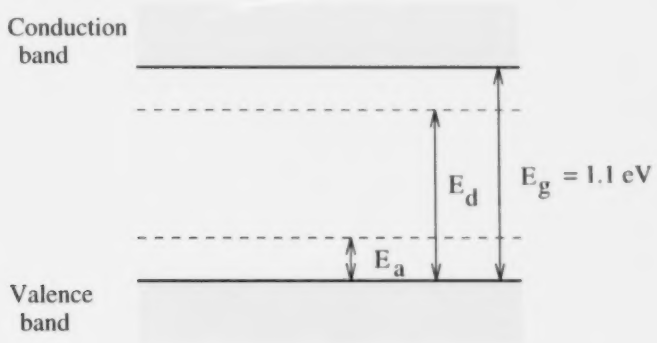


Figure 3-2. Energy levels created by doping intrinsic silicon.

For NMOS devices, the source and drain are commonly doped by diffusing phosphorus, arsenic, or antimony into the silicon at high temperature. This may be done by using a gas, liquid, or solid doping source. Two steps are required to dope the source and drain. The first step, known as predeposition, creates a local doping source on the surface of the silicon wafer. A first order approximation of the resulting vertical doping concentration, based on Fick's laws of diffusion, is given by

$$C(x,t_1) = C_{S1}\,\mathrm{erfc}\left(\frac{x}{2\sqrt{D_1 t_1}}\right) \qquad (3.4)$$

In the above equation, x is the diffusion depth, t_1 is the predeposition time, C_{S1} is the concentration of dopant at the silicon surface, D_1 is the dopant diffusivity, and erfc is the complementary error function. The solution above assumes that $C(0,t_1) = C_{S1}$ and $C(\infty,t_1) = 0$. The amount of dopant per unit area, Q, can also be defined as

$$Q = \frac{2}{\sqrt{\pi}} C_{S1} \sqrt{D_1 t_1} \tag{3.5}$$

The second step in diffusion doping, known as the drive-in, makes the dopant profile deeper and more uniform. If it is assumed that a constant amount of dopant exists, i.e. Q is constant, the concentration profile after drive-in can be approximated by

$$C(x,t_2) = \frac{Q}{\sqrt{\pi D_2 t_2}} \exp(-\frac{x^2}{4D_2 t_2}) \tag{3.6}$$

$$C(x,t_2) = C_{S2} \exp(-\frac{x^2}{4D_2 t_2}) \tag{3.7}$$

This requires removing the dopant source before drive-in begins. The solution above is of a Gaussian form.

3.3 Etching and Wet Chemical Cleaning

Film etching is done by a wet chemical or dry plasma etch. Wet chemical etching involves immersing a substrate into a heated or room temperature liquid etchant. Each wet etching step varies from the other, so the process details will be discussed in the appropriate sections. This section will provide some background on dry etching and wet chemical cleaning.

3.3.1 Reactive Ion Etching (RIE)

Reactive ion etching is a form of plasma etching in which etching proceeds by a combination of chemical and physical effects. By introducing an appropriate inert molecular gas into a glow discharge, reactive ionic species can be produced. These reactive species are directed by electric fields into the substrate surface where they chemically react to produce volatile products which can then be pumped away. It has been found that halogen gas ions are very effective for etching a large number of materials. Often, an additional gas is mixed with the halogen source to enhance etching rate or selectivity of etching one material over another.

The plasma in RIE etching is normally sustained using an RF power supply placed on the substrate electrode. The electrode develops a negative DC bias and positive ions are extracted from the plasma towards the substrate. The ions physically erode the substrate surface and provides more anisotropic etching than would be possible with ordinary plasma etching because of the directionality of the fields.

3.3.2 RCA Cleaning

Under high temperatures, contaminants on a wafer may migrate or diffuse and affect device performance. To prevent wafer contamination by molecular, ionic, and atomic contaminants, a wet chemical clean is usually done prior to any high temperature step. The two step "RCA clean", proposed by Kern and Puotinen, is most often used for this purpose [15]. The first step removes any light organic contamination. Organic contaminants can make a wafer surface hydrophobic to subsequent cleaning steps so these are removed first. The second step removes ionic and atomic contaminants. It is chosen to remove heavy metals and to prevent them from re-depositing onto the wafer.

The RCA clean procedure used for the NMOS process here has been previously investigated [16,17]. Wafers were first cleaned in a Pyrex beaker

containing a solution of 5:1:1 parts by volume H_2O:NH_4OH:H_2O_2. This was done for approximately 15 minutes at a temperature of about 80°C. Wafers were then rinsed in high resistivity deionized water. This was followed by a one minute etch in 100 parts H_2O and 1 part HF at room temperature. This is an additional step to the RCA clean and is intended to remove any native oxide from the wafer surface. Any underlying inorganic contaminants can then be removed more effectively by the last cleaning step. After another rinse, wafers were placed in an 80°C solution of $5H_2O$:$1HCl$:$1H_2O_2$ for 15 minutes.

Chemical solutions used in the clean were 27% by volume NH_4OH, 30% unstabilized H_2O_2, 37% HCl, and 49% HF. All were semiconductor grade. The deionized (DI) water used had a resistivity of greater than 16 MOhm-cm.

After cleaning, wafers were rinsed and spun dry in a Semitool spin rinser. Rinsing was done for two minutes at 800 rpm. Drying was done in heated nitrogen and lasted for one minute and 20 seconds at 2600 rpm. This was followed by a 40 second ramp down to 0 rpm in heated nitrogen. The nitrogen gas was from the Alberta Microelectronic Centre's in-house supply. This supply is boiled off from semiconductor grade liquid nitrogen. The gas must travel through piping some distance from the liquid nitrogen source, so its exact purity is not known.

3.4 NMOS Process Details

3.4.1 Mask Design

Prior to fabrication, a set of four masks were designed using Tanner Research's L-edit layout program. Mask design was done "in-house", but mask fabrication was done elsewhere by Precision Photomask. Mask design followed Mead-Conway design rules [18]. These rules specify the minimum allowable values for widths, separations, extensions, and overlaps for objects in the layout. The values are expressed in terms of a lambda value. In this design, a lambda value of 2.5 μm was used.

Devices and structures designed included a variety of different sized transistors, van der Pauw structures (Greek and offset quadrate crosses) for sheet resistance measurements of diffusions and polysilicon, contact strings and four terminal Kelvin test structures for determining contact resistance of diffusions and polysilicon, and passive devices like capacitors and resistors. Test structures and resistors were made from both diffused regions and doped polysilicon. Transistors were connected three in parallel (common source and gate) to simplify testing. Each transistor had different widths and lengths to confirm scaling.

To reduce prototyping costs, all four mask designs were combined onto a single mask substrate. The first mask was designed to be at the top left corner. Alignment with subsequent mask layers required successive 90° counter clockwise turns to line up properly with the underlying wafer. This procedure only produced working devices in the original top left quadrant of the wafer. A total of 16 prospective die per wafer were fabricated with this method. Each had a width of about 3 mm. Figure 3-3 gives a schematic representation of the 4 quadrant mask design and the overall mask layout for one die is shown in figure 3-4.

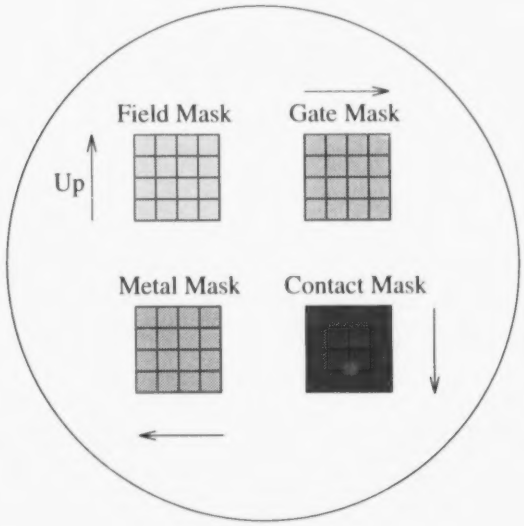


Figure 3-3. Four mask design on one mask substrate.

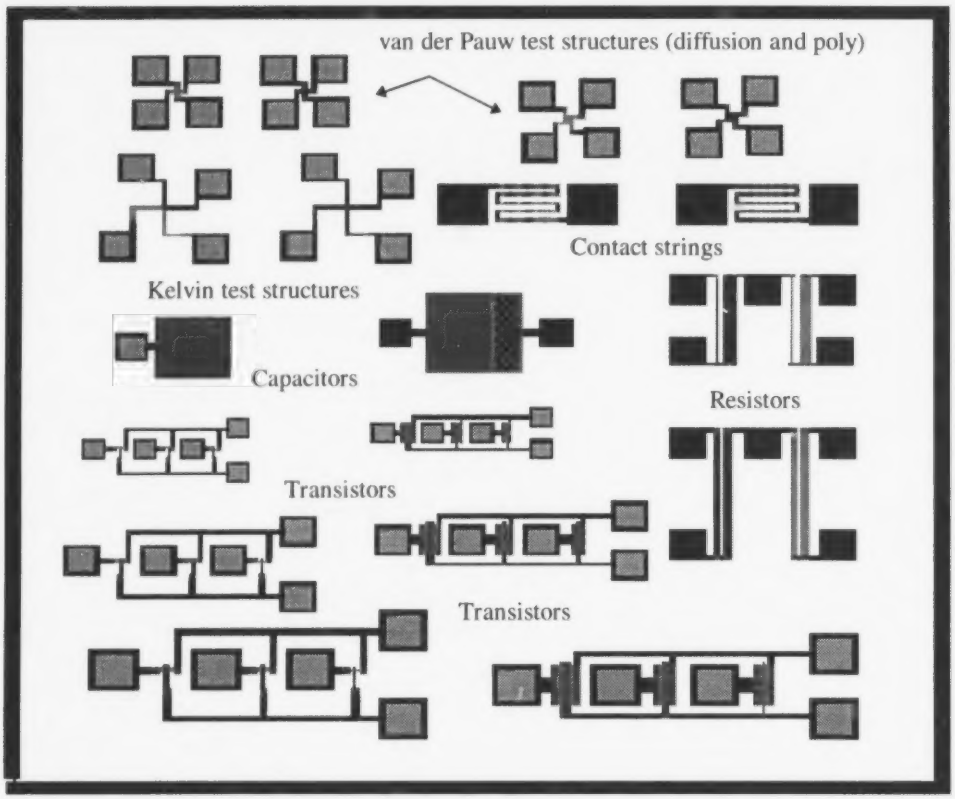


Figure 3-4. NMOS mask layout.

3.4.2 Starting Wafers

All devices and structures were fabricated on <100> orientation p-type (boron doped) test grade silicon wafers. <100> wafers have the lowest surface density of atoms when compared to <111> or <110> orientation wafers. This makes them

particularly suitable for MOS fabrication since surface impurities and trapped charge will have the least effect on device properties. Wafers were 3 inches in diameter and about 15 ± 2 mils (1 mil = 25.4 μm) thick, as specified by the manufacturer. Starting wafer resistivity ranged from 1 to 10 ohm-cm and was individually measured prior to fabrication.

3.4.3 Pad Oxidation

Oxidation was carried out in an 221 cm long horizontal tube oxidation furnace with an inner diameter of about 15 cm. Oxidation recipes were programmed into a Integrated Time Systems Tymkon process sequence controller. Furnace temperature was controlled with a 3 zone Tempress Model 261 digital temperature controller to create a large flat zone. Wafers were placed in a fused silica boat containing additional dummy wafers during oxidation. This boat was kept in a 600°C furnace tube at all times except during wafer loading and unloading. Wafer placement is shown in figure 3-5.

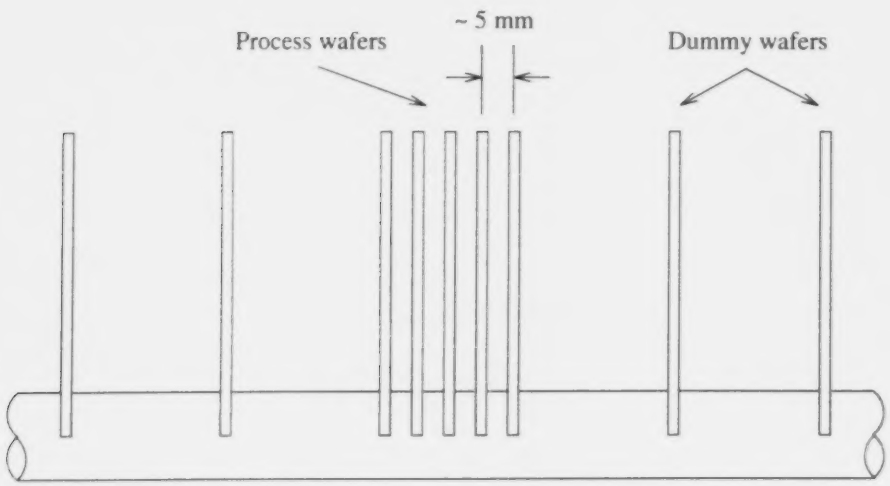


Figure 3-5. Wafer placement during oxidation.

Prior to actual oxidation, wafers were cleaned by the method described in section 3.3.2. The boat and furnace tube were also cleaned to reduce possible mobile ion contamination. The furnace tube used for oxidation was unlined and mobile ions are able to diffuse through the tube, or come from within the quartz tube itself, and contaminate the wafers inside. Sodium ions are the main source of mobile ion contamination and make up a large part of the charge, Q_m, discussed in section 3.1.1. By using chlorine and chlorine-containing compounds, it has been found that the mobile sodium ion concentration can be reduced and oxidation rate enhanced [19,20]. It is believed that sodium reacts with chlorine to form NaCl. NaCl is volatile at high temperatures and is swept away by the flowing gases in the furnace tube. Sources of

chlorine include pure Cl_2 gas and HCl gas. However, because of the hazards associated with chlorine gas and the corrosive nature of HCl gas, an alternative liquid source of chlorine involving TCA (1,1,1-Trichloroethane, $C_2H_3Cl_3$) was used.

The TCA delivery involved a flow of 99.993% pure oxygen and nitrogen carrier gas bubbled through 99.99995% liquid TCA which was kept in a quartz flask in a Schumacher bubbler held at 20°C. Approximate flow rates used were 1.5 litres/minute of oxygen and 75 sccm of nitrogen carrier gas. At 20°C, the N_2 carrier gas produced a flow of about 16 sccm of TCA. Cleaning consisted of a 45 minute flow of oxygen and TCA and then a 30 minute flow of only oxygen in a 1000°C tube. These times were suggested by the TCA manufacturer. The latter flow of oxygen allowed any remaining TCA time to fully react. TCA reacts with oxygen in the tube according to the following reaction

$$C_2H_3Cl_{3(g)} + 2O_{2(g)} \rightarrow 3HCl_{(g)} + 2CO_{2(g)} \qquad (3.8)$$

After cleaning the wafers and furnace tube, wafers were loaded into the quartz boat and slowly pushed into a 1000°C tube at a rate of 510 mm/min (20"/min). This was done under an oxygen flow of 1.5 ℓ/min. Wafer loading and unloading speed was controlled using a Tempress boat loader. Oxidation was then carried out by flowing 1.5 ℓ/min O_2 and 75 sccm N_2 carrier gas (TCA) for 60 minutes. This was followed by a 28 minute oxidation in 1.5 ℓ/min oxygen. The mechanism for enhanced oxide

growth in a chlorine ambient is not completely known, but published experimental data relating HCl concentration and oxide thickness was used to estimate oxide growth rate [21]. A conversion factor of 3% HCl = 1% TCA was assumed based on the reaction given above. The times used for oxidation were estimated from oxide thickness charts to obtain a target thickness of 80 nm [21,22]. Experimentation with the time for the final oxidation step was required to obtain this thickness. Following oxidation, the oxide thickness was verified using optical interferometry. In particular, a Nanometrics Nanospec/AFT film thickness measurement system was used. Measurements were taken at five points on the wafer and averaged. Values were found to be within 2% of each other.

Immediately following oxidation, wafers were annealed in the furnace tube (in situ) under a 1.5 ℓ/min nitrogen flow at process temperature for 20 minutes. Again, the AMC's in-house supply of nitrogen gas was used.. Annealing in an inert ambient such as nitrogen allows any unreacted silicon time to oxidize fully before being pulled from the furnace. This minimizes the number of ionized silicon sites and, hence, the fixed oxide charge, Q_f. Although fixed oxide charge is not a critical concern for the pad oxide, the annealing step was done as a precaution. After annealing, wafers were pulled from the furnace at a rate of 200 mm/min (8"/min) under 1.5 ℓ/min of nitrogen

and allowed to cool. Deal and others have shown that there is a reduction in fixed oxide charge after annealing and pulling in nitrogen [23, 24].

3.4.4 Silicon Nitride Deposition

After pad oxidation, nitride deposition was done using PECVD in a Plasma-Therm System VII 790 parallel plate reactor. In this reactor, process gases flow into the chamber through the top perforated aluminum electrode from above. After a chamber clean and pre-conditioning with the process gases, wafers were placed onto a substrate chuck (electrode) heated to 350°C. Gases and flow rates used were 2.5 sccm of NH_3, 6 sccm of SiH_4, and 280 sccm of N_2. Gas purities were 99.999%, 99.9%, and 99.999%, respectively. The primary reaction proceeds as [25]

$$SiH_{4(g)} + NH_3\,(N_2)_{(g)} \rightarrow Si_xN_yH_{z(s)} + H_{2(g)} \qquad (3.9)$$

The main source of nitrogen in the above reaction is from ammonia because it dissociates into atomic nitrogen more readily than does nitrogen gas. The main role of N_2 is to sweep out impurities and exhaust gases. It is also important to have an excess of nitrogen present to prevent silicon-rich film formation. The resulting silicon nitride product is nonstoichiometric and may contain from 10 to 35 atomic percent hydrogen [26]. Flowing nitrogen gas helps reduce this amount. This nitride

recipe was developed at the AMC from past experience and also included a 620 sccm flow of 99.999% pure helium gas to dilute the gas mixture. Deposition conditions are summarized below in table 3-1. An average nitride thickness of about 800 Å was obtained for the above conditions, as measured by a Nanospec/AFT system.

Parameter	Setting
Ammonia (NH_3)	2.5 sccm
Silane (SiH_4)	6 sccm
Helium (He)	620 sccm
Nitrogen (N_2)	280 sccm
RF Power (13.56 MHz)	55 Watts
Deposition Pressure	1 torr
Temperature	350°C
Time	7 min. 12 sec.

Table 3-1. PECVD silicon nitride deposition parameters.

3.4.5 Field Region Pattern Etch

The first mask used in the NMOS process defines the field oxide and active regions. This is done by patterning the silicon nitride and pad oxide. The first step is to coat the wafers with photoresist. To enhance the adhesion of the resist to the nitride, wafers were first vapour primed with hexamethyldisilazane (HMDS). HMDS is a commonly used adhesion promoter and ties up any water on a wafer surface [27]. Priming was done in a Yield Engineering Systems LP III-M5 oven set to 150°C.

After the application of HMDS and 10 to 15 minutes of cooling, wafers were spin coated with HPR 504 positive resist. This was done on a Solitec 1111-CTD photoresist spin coater. Settings were a 4 second 500 rpm spread followed by a 16 second 3500 rpm spin. These settings were found to be adequate for the device geometries on the designed mask. A resist thickness of about 1.7 micrometers was measured using a Nanospec/AFT immediately after spinning.

After resist spin, wafers were vacuum softbaked at 110°C in a Solitec BPM 2000 oven. This consisted of a 50 second vacuum bake, 10 second nitrogen fill under vacuum, and finally a 10 second nitrogen backfill. Softbaking removes almost all of the solvents in a photoresist, thereby making the resist photosensitive and ready for exposure to ultraviolet light.

Resist exposure was done using a Karl Suss MA4 contact mask aligner. Power level was set at a constant 350 W. Exposure times from the mercury bulb ranged from 2.45 to 4 seconds depending on the monitored bulb intensity at the time. Both standard (ST) and high precision (HP) modes on the Suss aligner were used during exposure. ST mode maintains mask-wafer contact by using a flow of nitrogen beneath the wafer. HP mode provides a vacuum between the mask and wafer during exposure. HP mode theoretically provides better precision, but no noticeable differences were detected after exposure.

After resist exposure, wafers were spray-developed in a Solitec 5110-CD developer system using 354 developer. Wafers were first pre-wet for 5 seconds at a speed of 500 rpm. This was followed by a 55 second develop at 460 rpm, a 90 second rinse at 500 rpm, and then a 40 second 1500 rpm spin dry. Times and rotation speeds were determined by past experience and some experimentation [28].

Developed wafers were then inspected by optical microscope for sharp features and bridging and those that did not turn out were reworked. To prepare the wafers for etching, they were hardbaked in a 140°C Despatch LCC1-54 oven for about 30 minutes. Hardbaking drives off any remaining solvents left after softbaking, dries the wafer of water, and removes resist pinholes by causing the resist to slightly flow. These effects make the photoresist more resilient during etching.

The patterned photoresist now acts as a mask for nitride etching. Etching was done in a parallel plate Plasma-Therm Batchtop RIE. The reactor chamber was first cleaned using 50 sccm of 99.9999% pure oxygen for 5-10 minutes, depending on contamination from previous RIE usage. RF power was set at 300 W and pressure set to 400 mtorr, as determined from past experience. Next, the chamber was pre-conditioned using the process gases. These were 95 sccm of 99.999% pure CF_4 and 5 sccm of oxygen. Again, this was done for 5 minutes, but at a power setting of 100 W and a pressure of 300 mtorr [29]. Then, the wafer was loaded and the nitride etched according to [11,30]

$$CF_{4(g)} \rightarrow 2F^{*} + CF_{2(g)} \tag{3.10}$$

$$Si_3N_{4(s)} + 12F^{*} \rightarrow 3SiF_{4(g)} + 2N_{2(g)} \tag{3.11}$$

Freon-14 (CF_4) in the plasma produces free radical fluorine. This fluorine reacts with the silicon nitride to form volatile silicon tetrafluoride. The addition of a small percentage of oxygen increases the number of fluorine radicals. This occurs by the reaction

$$CF_{4(g)} + O \rightarrow COF_{2(g)} + 2F^{*} \tag{3.12}$$

The extra fluorine contributes to enhance the etch rate. Silicon dioxide is also etched by CF_4/O_2, but silicon nitride etches faster because the Si-N bond is intrinsically weaker [30]. The selectivity of silicon nitride over silicon dioxide allows one to etch the nitride without losing too much of the underlying silicon dioxide. Before actual wafer etching, an estimate of the etch time was determined using test wafer pieces. Actual wafer etch times were found to be about 1 minute. In this time, all the nitride was etched, and roughly 120 Å of oxide was etched.

Before etching the remainder of the pad oxide, the photoresist was stripped. The patterned nitride underneath was then used as an etch mask. Resist removal was done by ultrasonically cleaning the wafers in acetone for about a half hour and then putting them into a so-called "piranha" bath for about 30 minutes. The piranha bath consisted of 3 parts concentrated sulfuric acid and 1 part 30% hydrogen peroxide.

Both were of semiconductor grade. The piranha bath removed any remaining resist by converting it to carbon dioxide and water and removed any film residue that may have been left by the acetone clean.

Pad oxide removal was done with a commercially available buffered oxide etch (BOE). This etchant is composed of ammonium fluoride and hydrofluoric acid. A semiconductor grade solution with a ratio of 10:1 NH_4F:HF was used. An etch time of 1 minute 20 seconds completely removed the oxide. The time used was based on earlier experiments which produced etch rates of about 560 Å/minute. Slight overetching was not a prime concern since device geometries were large compared to film thickness to be etched. Etch completion was detected by noting if the wafer surface was hydrophobic. A bare silicon surface is hydrophobic whereas one still covered with an oxide layer is not.

3.4.6 Field Oxidation

As previously mentioned in chapter 2, silicon nitride oxidizes very slowly in an oxygen or steam ambient. The patterned nitride on oxide from the previous step now acts as a convenient mask for field oxidation of the silicon wafer. Since device isolation requires a fairly thick layer of oxide, this step involves the faster wet

oxidation. The system at the Alberta Microelectronic Centre uses pyrogenic steam. High purity hydrogen and oxygen are mixed in the furnace tube at a temperature higher than the H_2/O_2 combustion temperature of 620°C. To prevent possible explosion, the system is set up with interlocks preventing H_2/O_2 flow below 800°C and at a ratio of more than 2:1 H_2:O_2.

The oxidation process began with wafer and furnace cleaning according to section 3.4.3. This time the oxidation temperature of 1100°C was used for cleaning. Wafer loading was done at 1000°C under 1.5 ℓ/min oxygen with a loader speed of 510 mm/min. The furnace was then ramped up to 1100°C at a rate of 10°C/min. This took approximately 10 minutes. Immediately after reaching 1100°C, the process recipe was started [29]. This began with a 10 minute dry oxidation in 1.5 ℓ/min oxygen. Following that was a wet oxidation in 1.0 ℓ/min 99.999% pure hydrogen and 0.6 ℓ/min oxygen for 3 hours. Problems later encountered with the mass flow controller for hydrogen required the recipe to be changed to a 0.9 ℓ/min flow of hydrogen and 1.0 ℓ/min flow of oxygen. This step was followed by a 30 minute oxidation in 1.5 ℓ/min O_2 and 16 sccm TCA and then by a 10 minute dry oxidation in 1.5 ℓ/min oxygen. Annealing was then performed for 40 minutes under 1.5 ℓ/min N_2. With about 10 minutes remaining in the anneal, the furnace temperature was ramped

down to 1000°C and the wafers pulled out at a rate of 203 mm/min under nitrogen flow. Oxide thickness had an average value of about 9200 Å when measured with a Nanospec/AFT. This compares with a calculated value of 1.2 μm for these conditions using equation 3.3.

It is of interest to mention that the growing field oxide diffuses slightly under the nitride mask during oxidation. The grown field oxide then takes on a tapered appearance, resembling a bird's beak, along the nitride edges. A scanning electron microscope photograph of this phenomenon is shown in figure 3-6. The top film in the figure is a PECVD deposited layer of silicon dioxide.

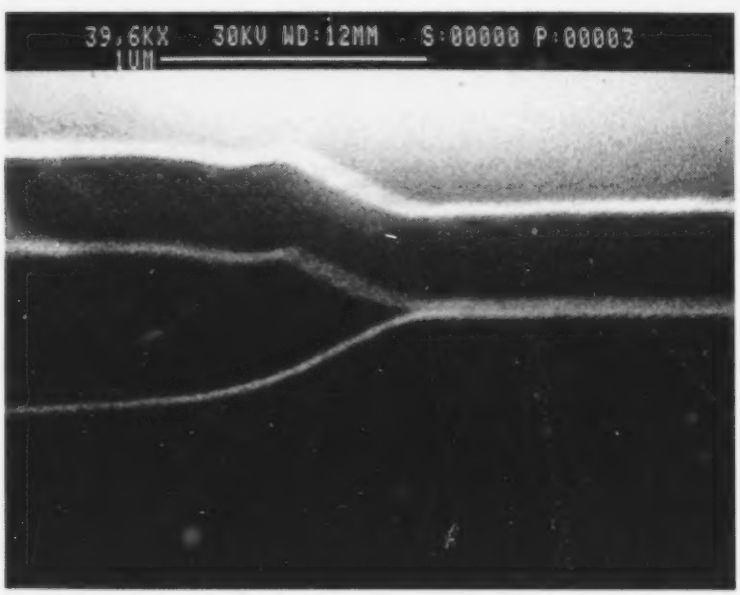


Figure 3-6. Bird's beak formation in field oxidation.

3.4.7 Nitride and Pad Oxide Removal

The slow oxidation of silicon nitride in a dry and wet oxidation ambient occurs by the following reactions [31]

$$Si_3N_{4(s)} + 3O_{2(g)} \rightarrow 3SiO_{2(s)} + 2N_{2(g)} \qquad (3.13)$$

$$Si_3N_{4(s)} + 6H_2O_{(g)} \rightarrow 3SiO_{2(s)} + 6H_{2(g)} + 2N_{2(g)} \qquad (3.14)$$

Because of the relatively short times used in field oxidation in a dry ambient, nitride conversion to oxide was predicted assuming a solely wet ambient. Using data published by Fränz and Langheinrich, about 650 Å of SiO_2 was predicted to be formed [31]. This converted nitride was etched with a solution of 360 mL DI water and 40 mL 49% HF before stripping the remaining nitride. Previous experiments on sample wafer pieces with this etchant produced an etch rate of about 430 Å/min. Process wafers were etched for a total of 2 minutes in this solution to completely remove this oxide layer. This left an average nitride thickness of about 290 Å when measured with a Nanospec/AFT. The remaining nitride was then etched using RIE for approximately 45 seconds using the conditions previously used in section 3.4.5. An average remaining pad oxide thickness of about 650 Å was left. This oxide layer was initially removed using a 10:1 BOE etchant for about 1 min 30 seconds. This was found to etch more of the field oxide than desired so a time of 1 min. 10 seconds was subsequently used.

3.4.8 Gate Oxidation

Two different gate oxide thicknesses were investigated for the NMOS process. One was 500 Å thick and the other was 800 Å. The thinner oxide was expected to result in lower threshold voltages and higher transconductance, but be more difficult to control and reproduce. Both oxide layers were grown at 1000°C in a dry oxidation ambient, but the 800 Å oxidation also included TCA flow which enhances oxidation growth. Before oxidation, the furnace tube and wafers were cleaned according to section 3.4.3. Wafers were then pushed into a 1000°C tube at a rate of 510 mm/min under 1.5 ℓ/min oxygen. 500 Å oxides were grown for 60 minutes in 1.5 ℓ/min oxygen. 800 Å oxides went through a 10 minute dry oxidation in 1.5 ℓ/min O_2, a 60 minute oxidation in 1.5 ℓ/min O_2 and 16 sccm TCA, and then a 28 minute dry oxidation in 1.5 ℓ/min O_2. Following oxidation, all wafers were annealed for 40 minutes in a 1.5 ℓ/min nitrogen ambient to reduce fixed charge which is critical to the gate oxide. Wafers were then pulled from the furnace at a rate of 200 mm/min in the same nitrogen flow rate.

3.4.9 Polysilicon Gate Deposition

As with the nitride, polysilicon was deposited using PECVD. The reactor chamber was again cleaned and pre-conditioned prior to deposition. The deposition of polysilicon was done by pyrolyzing silane gas. The reaction is given by [26]

$$SiH_{4(g)} \rightarrow Si_{(s)} + 2H_{2(g)} \qquad (3.15)$$

A 50 sccm flow of 99.9% pure silane was used in the deposition. Deposition conditions were based upon a recipe suggested by the PECVD reactor manufacturer.

Polysilicon deposited by PECVD is amorphous and becomes polycrystalline during subsequent high temperature steps. Films deposited at 350°C showed signs of stress and portions did not adhere well to the wafer surface. It was later found that films deposited at a lower temperature of 250°C did not peel and blister as much as those deposited at 350°C [32]. A thickness of about 0.45 microns was obtained with the process conditions summarized in table 3-2 below.

Parameter	Setting
Silane (SiH_4)	50 sccm
RF Power (13.56 MHz)	25 Watts
Deposition Pressure	1 torr
Temperature	350°C or 250 °C
Time	15 min. 30 sec.

Table 3-2. PECVD polysilicon deposition parameters.

3.4.10 Gate Pattern Etch

Polysilicon was patterned using a positive resist mask. The resist was patterned using the same method described in section 3.4.5. After resist hardbake, the polysilicon was etched by reactive ion etching. This was mainly due to difficulties in finding a suitable wet etchant for polysilicon which did not also etch the photoresist mask. In this step, the RIE chamber was again cleaned and conditioned with the same gases mentioned in section 3.4.5. Free radical fluorine, in this case, etches the silicon by forming SiF_4 according to the reaction [30]

$$Si_{(s)} + 4F^{*} \rightarrow SiF_{4(g)} \qquad (3.16)$$

The CF_4/O_2 chemistry preferentially etches silicon faster than silicon nitride and silicon dioxide. Complete removal of the polysilicon was done in only 3.5 minutes. This etch time only removed about 50 Å of the underlying gate oxide for both 500 Å and 800 Å gate oxide wafers. Following gate patterning, the photoresist mask was removed using an ultrasonic acetone bath and a piranha bath for about 30 minutes each.

3.4.11 Source/Drain Gate Oxide Etchback

The exposed gate oxide layer underneath the polysilicon was removed using 10:1 BOE etchant. For wafers with 500 Å gate oxides, the etch took about 50 seconds. For 800 Å gate oxides, the time was about 1 min 10 sec. Etch completion was verified by obtaining a hydrophobic wafer surface and with thickness measurements.

3.4.12 Source/Drain Diffusion

N-type doping of the source and drain was done using the self aligned process. The patterned polysilicon layer was used as a mask for source/drain diffusion, and the thick field oxide regions prevented doping of the silicon between devices. To sufficiently mask dopant diffusion, the field oxide must be thicker than the diffusion length of the dopant used in predeposition. For the conditions described here, a minimum field oxide thickness of about 0.1 μm was required to mask doping [33]. Measured field oxide thickness was at least 5 times greater than the required amount.

Predeposition doping was done with 3" phosphorus solid doping sources in a diffusion furnace identical to the one used for oxidation. In this method, wafers and

doping sources are loaded into the furnace under high temperatures. Dopant vapour then diffuses from source wafers and reacts with adjacent silicon wafers. A phosphosilicate glass (PSG) layer forms on the wafers and elemental phosphorus then diffuses into the silicon. The solid doping sources used in this process (PH-900) were supplied by Carborundum. The PH-900 n-type phosphorus planar diffusion source consists of cerium pentaphosphate carried on and in an inert substrate. At diffusion temperatures, the CeP_5O_{14} decomposes to form P_2O_5 vapour. This is given by the reaction

$$CeP_5O_{14(s)} \rightarrow Ce(PO_3)_{3(g)} + P_2O_{5(g)} \qquad (3.17)$$

The transport of P_2O_5 vapour to process wafers does not require a carrier gas for mass transport. However, a flow of an inert gas is usually used to prevent back flow of contaminants into the diffusion tube and to provide a boundary layer around the doping sources and wafers. Monkowski and Stach have found that doping uniformity is affected by the gas flow rate of the inert gas and by wafer spacing in the boat [34]. Dopant source and wafer spacing for the predeposition step is shown in figure 3-7. Process wafer front surfaces face the adjacent source. Dummy wafers were used to control and standardize gas flow.

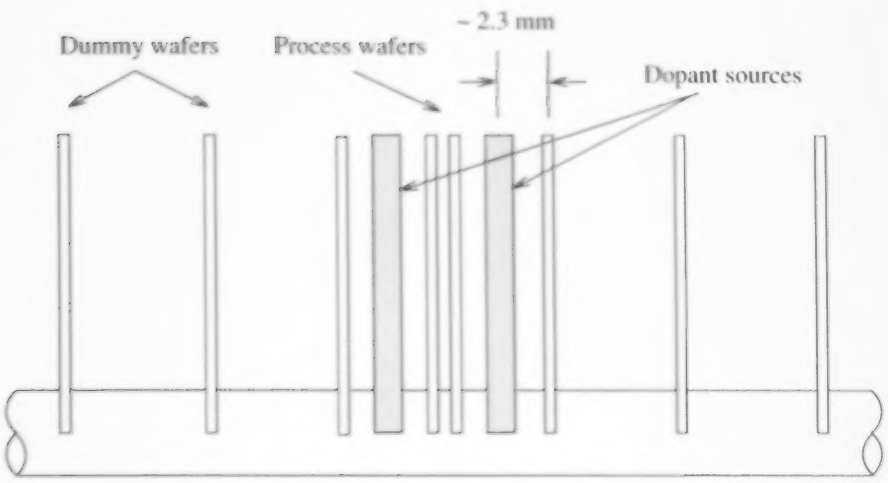


Figure 3-7. Wafer and doping source placement during predeposition.

Prior to use, new doping sources were annealed at 925°C for 16 hours under 4.0 ℓ/min N_2. Between uses, the sources were stored in the diffusion tube under 1.5 ℓ/min N_2 at a temperature of 600°C. If dopant wafers were not used for over a month, they were reannealed at 925°C for 8 hours under 2.0 ℓ/min N_2.

Predeposition conditions involved a 40 minute diffusion at 860°C. Wafers were first RCA cleaned, placed into the wafer boat with the doping sources, and then pushed into the furnace at a loader speed of 430 mm/min. This was done at a temperature of 860°C and under a flow of 1.5 ℓ/min nitrogen. Initially during diffusion, a boundary layer around the wafers was formed with a 4.0 ℓ/min flow of N_2.

This flow rate caused wafers to be more highly doped at their centres. A possible cause of the non-uniform wafer doping was due to turbulence around the wafer periphery. Subsequent predepositions using a 2.0 ℓ/min flow of nitrogen did not produce as much turbulence and gave more uniform doping. After 40 minutes, wafers were pulled from the furnace under the same conditions as loading.

Once wafers were cool, the surface PSG layer was etched off using a 360 mL H_2O and 40 mL 49% HF solution. This step was done to allow drive-in to occur without further doping and to prevent phosphorus contamination of the drive-in furnace tube. Only a 15 second etch was required to completely remove the PSG. Etch completion was signaled by a hydrophobic silicon wafer surface.

Following PSG removal, drive-in diffusion was performed in a 147 cm long diffusion tube with an inner diameter of about 15 cm. Wafers were pushed and pulled at a rate of 357 mm/min under 1.5 ℓ/min O_2. Wafer placement in the boat was identical to the arrangement used for oxidation. Drive-in lasted for 30 minutes under 1.5 ℓ/min O_2 at a temperature of 1050°C. A drive-in diffusion done in an oxygen ambient enhances the diffusivity of phosphorus and also regrows a thin oxide layer on any silicon surfaces [35]. Immediately following drive-in, the furnace temperature was ramped down to 600°C as the wafers were being pulled out.

3.4.13 Dielectric Deposition

The deposition of a dielectric layer after drive-in allows isolation between the metal layer and the underlying silicon. This dielectric layer also provides better device isolation as it adds to the thickness of the existing field oxide and prevents any undesirable conduction between active regions. The process described here used a PECVD deposited layer of SiO_2 as the dielectric. By reacting silane and nitrous oxide in a plasma, it is possible to deposit SiO_2 by the reaction

$$SiH_{4(g)} + 2N_2O_{(g)} \rightarrow SiO_{2(s)} + 2H_{2(g)} + 2N_{2(g)} \qquad (3.18)$$

Nitrous oxide is a more effective source of atomic O than is O_2 due to the stronger bonding of the diatomic molecule. The process recipe used here was previously developed at the AMC and also included helium and nitrogen gas. Purity of the nitrous oxide was 99.999%. Silane, helium, and nitrogen gas purities were the same as those mentioned previously. Deposition parameters shown in table 3-3 gave a SiO_2 layer approximately 0.58 μm thick.

Parameter	Setting
Silane (SiH_4)	5.5 sccm
Helium (He)	600 sccm
Nitrogen (N_2)	100 sccm
Nitrous Oxide (N_2O)	500 sccm
RF Power (13.56 MHz)	25 Watts
Deposition Pressure	1 torr
Temperature	350°C
Time	14 min

Table 3-3. PECVD silicon dioxide deposition parameters.

3.4.14 Contact Pattern Etch

Patterning of contact holes in the deposited CVD oxide was done using a negative chrome mask and positive photoresist. With this arrangement, UV light only reaches the wafer surface where contact holes are desired. The rest of the wafer is masked and the underlying resist later serves as an etch mask. Positive photoresist patterning was done according to the procedure previously used for silicon nitride and polysilicon patterning. Etching of contact holes was done using 10:1 BOE with a surfactant to improve etchant wetting in the small contact holes. An initial etch time of 12 minutes was used. This was determined using test wafer pieces. Using this time, about 7500 Å of oxide was removed from actual process wafers, which was slightly higher than desired. Subsequent batches of wafers, with similar oxide thicknesses, only required a 6 to 7 minute etch to completely etch the contact holes.

These etches removed approximately 6200-6600 Å of oxide. The etch rate discrepancy between similar batches of wafers is unexplained since the same 10:1 BOE etchant was used. After etching, the photoresist was removed with a 30 minute ultrasonic acetone clean and a 30 minute piranha bath.

3.4.15 Front Metallization

Device interconnects on top of the wafer were made using aluminum metal. This aluminum was sputter deposited using a Perkin Elmer 4410 sputtering system. The 4410 system is a DC planar magnetron system which uses delta shaped targets and a rotating substrate holder. A 99.999% pure aluminum target was used for depositions. Chamber base pressure was pumped down to the low 10^{-7} torr range before all depositions.

Immediately prior to loading into the Perkin Elmer system, wafers were dipped in BOE. The purpose was to remove any native oxide from the contact holes. A 10 second etch in 10:1 BOE with surfactant was used for all wafers, but this step was unfortunately neglected for the first full process wafer. Deposition process parameters are shown in table 3-4. This process gave film thicknesses of about 1.1 μm, as measured by a Tencor Instruments stylus profilometer.

Parameter	Setting
Deposition Pressure	10 mtorr
Substrate to Target Distance	5.5 cm
Substrate Rotation	4.73 rpm
Power	2 kW
Target Voltage	-380 V
Deposition Time	30 min.

Table 3-4. DC planar magnetron aluminum sputter conditions.

3.4.16 Metal Pattern Etch

As before, the patterning of aluminum began with a patterned photoresist etch mask. Photoresist was patterned as described in section 3.4.5, but vapour priming with HMDS was not required for this step as adhesion to Al was already acceptable. Hard baking was initially done at 140°C, but later changed to 120°C because such a high temperature was not required [36]. Aluminum etching was done in a 50°C solution of Aluminum Etch III manufactured by KTI Chemicals Inc. This solution is composed of 90% phosphoric acid, 5% nitric acid, and 5% water. Etching was done in a Pyrex beaker on a hotplate/stirrer set to a speed of about 300 rpm. Aluminum was very easily etched in this solution, so overetching was a prime concern. It was often necessary to remove wafers from the etchant, rinse, and then inspect visually for etch completion. Wafers which were not completely etched were then blown dry in

nitrogen and placed back into the solution. Etch times varied from about 2 min to 3 min.

3.4.17 Backside Oxide Removal

To make a good substrate contact, the residual oxide on the back of wafers was removed. To prevent etching of the front side of wafers, black tape manufactured by Semiconductor Equipment Corp. was used to cover the entire front wafer surface. According to the manufacturer, this tape is a non-contaminating plastic film with a high tack adhesive on the back. The tape was trimmed and applied to the wafers and backside oxide was then removed using a cotton swab and 49% HF. Oxide thickness was on the order of the thickness grown during field oxidation. After etching, wafers were then rinsed and the black tape removed. No front side damage from the tape or etch was observed.

Hardbaked photoresist from the aluminum patterning step was still present on the wafer fronts so this had to be removed. First, wafers were placed into an ultrasonic acetone bath for about 30 min. To ensure complete resist removal, wafers were then rinsed, dried, and placed into a solution of Microstrip 2001. This solution is a positive resist stripper manufactured by Olin Hunt. Stripping was done at 90°C,

as suggested by Olin Hunt, in a Pyrex beaker on a hotplate/stirrer set for low agitation. This lasted for about 10 min and with an agitation speed of 100 rpm.

3.4.18 Metal Anneal

After forming Al/Si contacts, it is standard practice to anneal them in nitrogen at temperatures below the Al-Si eutectic point of 577°C. This "sintering" step has been found to reduce the contact resistance between the Al and Si [37]. For contacts between n-type silicon and aluminum, it has been found that sintering between 400°C and 450°C is the optimum range [37]. The sintering step also provides a convenient opportunity to reduce the number of interface states produced during oxidation. By using nitrogen gas with a small amount of hydrogen (forming gas), it is possible to reduce contact resistance and interface state density at the same time.

The process detailed here used two anneal temperatures. Four hundred degrees Celsius was used for the first batch of wafers and a temperature of 450°C was used for the others. Sintering was done for 20 minutes in a Thermco MB-80 "Mini-Brute" furnace using a forming gas composed of 8.12% hydrogen and the remainder nitrogen. The "Mini-Brute" uses 3 analog type controllers to set temperatures for 3 temperature zones across the furnace tube. The center zone is adjustable from 400-

1400°C and the front and back zones are adjusted ±50°C relative to the center zone setting. Dial settings for the center zone range from 000 to 999. The former corresponds to 400°C and the latter to 1399°C. Front and back settings also range from 000 to 999, but here a setting of 000 means a temperature of -50°C relative to the center setting. A zone setting of 999 gives a temperature 50°C hotter than the center zone. The abovementioned settings are nominal and actual furnace temperatures had to be monitored with a thermocouple prior to sintering. To obtain a desired temperature, furnace profiling was done for the center zone to determine the exact setting required for the center control. A more accurate representation of sintering conditions would have been to profile with forming gas flowing. However, with the first batch of wafers sintered at 400°C, this procedure was omitted and the furnace was profiled with no gases flowing. A middle setting of 20 was required to obtain a temperature of about 400°C. Later, a temperature of 450°C was profiled using a flow of 0.7 ℓ/min house nitrogen. Here, a setting of 85 was used to reach 450°C. Furnace temperature took a significant time to settle, so, to save forming gas, only nitrogen was used to profile. It was assumed that the temperature difference between using forming gas and an equivalent flow of pure nitrogen was insignificant. After determining the exact temperature setting and allowing the tube temperature to settle, a 0.7 ℓ/min flow of forming gas was purged through the tube for about 5 minutes. Wafers were then placed into a quartz boat and slowly pushed into the

furnace by hand. After 20 minutes, wafers were slowly pulled and allowed to cool. A front and middle setting of 650 and 950 was used for all anneals to provide a uniform furnace temperature. These values were determined by past experience with the furnace.

3.4.19 Back Metallization

The final step in the fabrication process was to deposit a conductive layer on the backside of wafers. This layer would provide good electrical contact for substrate connections and would be required for packaging of wafer die. Again, a sputtered aluminum film was used for this purpose. The same process for front metallization was used here, but with a shorter deposition time. Approximately 0.5 μm of aluminum was deposited with a 15 minute deposition.

4. Characterization

Characterization of the NMOS process consisted of measurements made at both wafer and device level. These measurements were taken during and after the fabrication process. Test wafers were used at various points in the fabrication to monitor process conditions. Test results at wafer and device levels will be presented in the following chapter. This chapter will discuss characterization techniques and equipment used for measurements.

4.1 Wafer Level Measurements

4.1.1 Nanospec/AFT Film Thickness Measurement System

The thicknesses of grown or deposited transparent films of silicon dioxide and silicon nitride were measured using a Nanometrics Nanospec/AFT film thickness measurement system. The Nanospec/AFT measures film thickness using a method of optical interferometry [33]. The system uses a lamp to illuminate the film to be measured. Interference from multiple reflections within the film forms an approximate sinusoidal variation as a function of wavelength. The frequency of the

sinusoidal variation is a function of film thickness and refractive index. The refractive index is entered by the user, so the system calculates a value for the film thickness. Wavelength is scanned from 480 nm to 800 nm and the generated spectrum is compared to a bare silicon reference. The Nanospec/AFT can measure silicon oxides on a silicon substrate in the range 400-30000 Å ± 2%, nitrides from 400-10000 Å thick, and positive resist from 500-40000 Å thick [38]. Refractive indices used for measurement of the above films were 1.45, 2.00, and 1.64, respectively. Film measurements consisted of entering a film type, performing a bare silicon reference scan, entering a film refractive index, and then taking sample measurements.

4.1.2 Tencor Instruments Alpha-Step 200

Opaque metal and polysilicon film thicknesses were measured using an Alpha-Step 200 profilometer manufactured by Tencor Instruments. A profilometer measures thickness with a diamond stylus which contacts the sample surface. The stylus is scanned across the sample and vertical deflections are converted to electrical signals. The profilometer then displays a scaled readout of height versus scan position. The Alpha-Step 200 has a vertical and horizontal resolution of 5 Å and 400 Å, respectively, when set to display vertical range in kiloangstroms [39]. To measure

the thickness of aluminum and polysilicon films, part of the substrate was masked during deposition to create a step. For Al films, a microscope cover slip was used to partially cover test samples during sputter deposition. Test wafers of PECVD deposited polysilicon were partially masked using pieces of silicon wafer. Alpha-Step scanning of test films began and ended at considerable distances from the masked film edge. This was done to minimize the measurement error due to film deposition shadowing caused by the masking material. Film thickness measurements taken too close to the interface would have given false results.

4.1.3 Four Point Probe

Measurements of wafer sheet resistance were done using a Veeco Instruments FPP-5000 four point probe. The usual four point probe configuration measures resistivity by using four collinear probes spaced an equal distance apart. Current is forced between the two outer probes and the voltage difference across the two inner probes is used to calculate resistivity. This arrangement minimizes the contribution of contact resistance and spreading resistance (resistance encountered by the current flowing out of the probe tip into the semiconductor) to the total resistance measurement because very little current flows in the inner probes. The low current produces very small voltage drops on the inner probes and contact and spreading

resistance become negligible. When compared to the four point probe inter-spacing, s, all sample thicknesses, t, were considered "thin" ($t \leq s/2$) and the equation for sheet resistance under such conditions is given by [40]

$$R_s = 4.53\left(\frac{V}{I}\right) \quad (\Omega/\text{sq.}) \tag{4.1}$$

Further correction factors are sometimes needed to compensate for wafer size and probe orientation and proximity to the wafer edge. Probe spacing on the Veeco four point probe was 1.59 mm (62.5 mils). For sheet resistance measurements taken on 76 mm (3 in.) wafers, as shown in figure 4-1, no additional correction factors were required [40]. In the figure, d is the distance to the wafer boundary and s is the probe spacing.

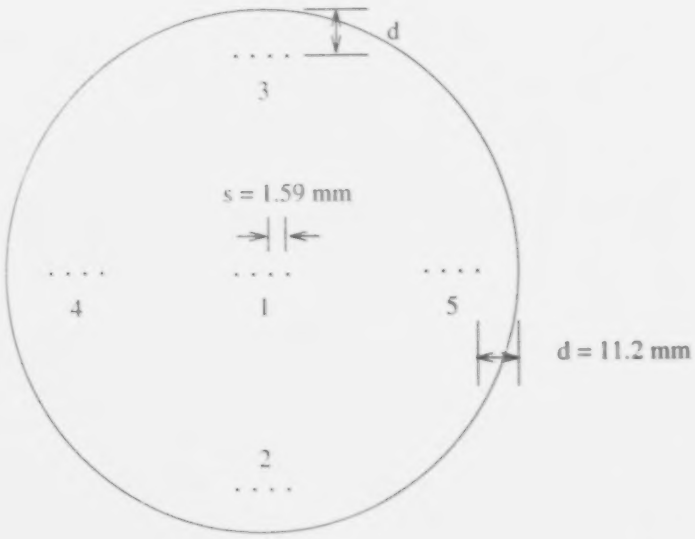
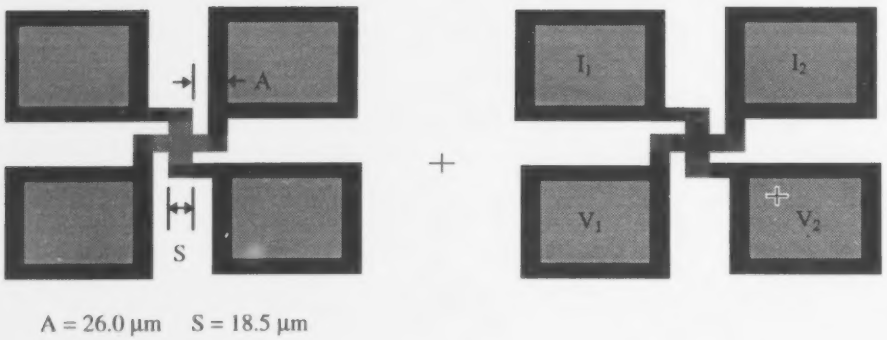


Figure 4-1. Four point probe placement for sheet resistance measurements.

Sheet resistance measurements were taken on new wafers prior to processing and on diffusion test wafers. The Veeco four point probe is also capable of determining semiconductor doping type (n or p) and this option was used in measurements [41].

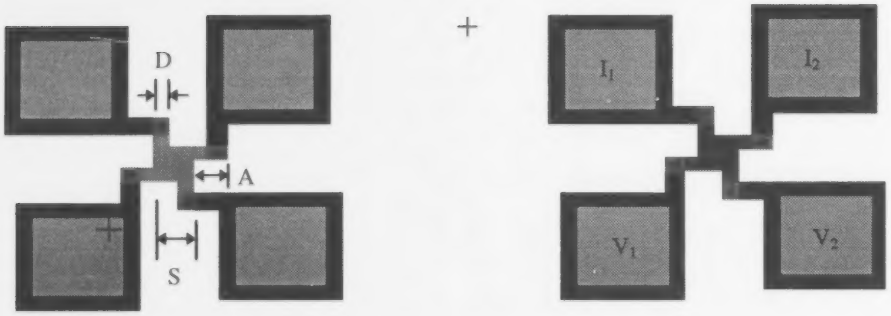
4.1.4 Van der Pauw Measurements

To measure sheet resistance of fully processed wafers, a method other than the usual four point probe had to be used. The van der Pauw technique allows sheet resistance to be measured using samples of arbitrary shape. This has the advantage of only requiring a small localized area for measurements. Symmetrical shapes are often used in van der Pauw measurements. In this work, two types of crosses were fabricated to measure sheet resistance. These are shown in figure 4-2.



A = 26.0 μm S = 18.5 μm

(a)



D = 15.0 μm A = 30.0 μm S = 37.5 μm

(b)

Figure 4-2. Van der Pauw test structures. (a) Diffused & polysilicon Greek crosses, (b) Diffused & polysilicon offset quadrate crosses. A=length of arm, S=width of central square, and D=width of arm.

As in four point probe measurements, the van der Pauw method uses a voltage to current ratio to calculate sheet resistance. The use of symmetrical shapes simplifies the sheet resistance equation to [40]

$$R_s = 4.53R\ (\Omega/\text{sq.}) \qquad (4.2)$$

where R is V_1-V_2 divided by the current from I_1 to I_2.

Using data published by David and Buehler [42] for the Greek cross A/S ratio shown above, a normalized sheet resistance error of less than 0.1% can be obtained. For the offset quadrate cross, the values for A, D, and S given above also produce a normalized error of less than 0.1%.

4.1.5 Diffusion Junction Depth

To obtain an estimate of source/drain doping concentration, it was necessary to determine the junction depth after drive-in diffusion. To delineate the p and n-type regions on a cross section of a diffused sample, a copper stain was used. A solution of 8 g/litre copper sulfate and 10 cm^3/litre of concentrated hydrofluoric acid (49% HF) was used to stain junctions [43,44,45]. Test samples were cleaved by hand from wafers and etched in 10:1 BOE for 15 seconds. Next, samples were rinsed, dried, and immersed in a 60 mL polypropylene bottle of the copper solution. This

was done for approximately 5 seconds under illumination from a ultraviolet lamp. Samples were then rinsed with deionized water and dried. Cross sectional views of diffused junctions were then obtained with a scanning electron microscope (SEM).

4.1.6 Contact Resistance

For the devices fabricated in this work, it is desirable to obtain ohmic behaviour at metal - semiconductor contacts. Ohmic contacts exhibit linear or quasi-linear current-voltage characteristics. Contact behaviour does not have to be strictly linear, but contacts should be able to supply enough current to the device. Also, the voltage drops across contacts should be small when compared to voltage drops across active device regions. To characterize contact formation, one can measure the contact resistance. A measurement of contact resistance usually includes not only the actual interface but also the regions immediately above and below the interface.

4.1.6.1 Contact Strings

The use of contact strings provides a rough estimate of contact resistance. A contact string is a series chain of contacts. A current, I, is forced from one end of the string and the resulting voltage, V, is measured. The total resistance is given by V/I.

Individual contact resistance can then be calculated by subtracting bulk semiconductor resistance from the total resistance and dividing by the number of contacts. Figures 4-3 and 4-4 show top and cross sectional views of the contact strings used for measurements. The structure is made up of twenty separate diffused resistors and forty 7.5 μm square contacts.

Figure 4-3. Top view of diffused and polysilicon contact strings.

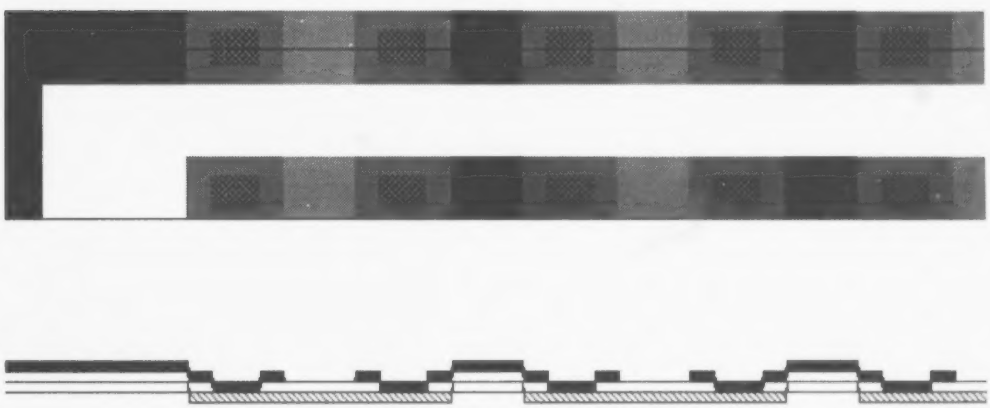


Figure 4-4. Cross sectional view of diffused and polysilicon contact strings.

4.1.6.2 Four Terminal Kelvin Test Structure

A more accurate measurement of contact resistance can be obtained by using a Kelvin test structure [46]. Contact resistance measurements using the four terminal Kelvin structure, or cross-bridge Kelvin resistor (CBKR), minimizes the parasitic effect from sheet resistance of the underlying semiconductor. As shown in figure 4-5, contact resistance is calculated by flowing a current between two contact pads, 1 and 2, and measuring the voltage difference between pads 3 and 4. Theoretically, contact resistance, R_c, is given by V_3-V_4/I. Specific contact resistance, ρ_c, can be calculated by $\rho_c = R_c A_c$, where A_c is the area of the contact. This normalized value can then be used to compare contacts of varying sizes.

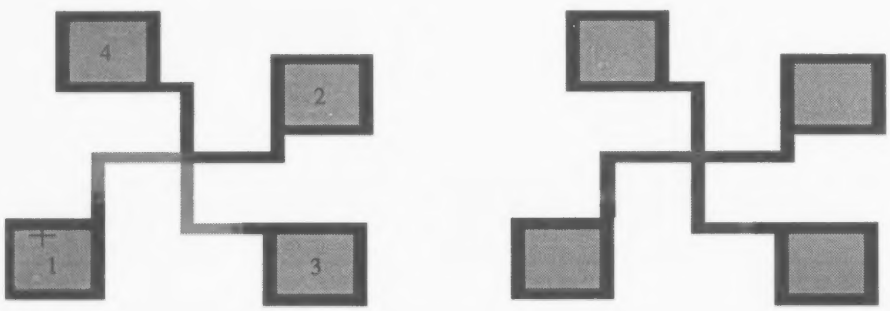


Figure 4-5. Diffused and polysilicon four terminal Kelvin test structures. Metal connects pads 2 and 4.

4.2 Device Measurements

4.2.1 Diffused and Polysilicon Resistors

The resistance of a doped region can be given by $R = R_s(L/W)$, where R_s is the sheet resistance and L and W are the length and width of the diffused region. Doping concentration and temperature both have direct effects on measured resistance. Doping is directly proportional to resistivity and temperature affects carrier mobility which is also proportional to resistivity. Measuring resistance of fabricated resistors simply consisted of ramping a voltage across the resistor and then plotting the resulting current at room temperature.

4.2.2 NMOS Transistors

4.2.2.1 DC Characteristics (I_D-V_{DS} Curves)

The dc characteristics of NMOS transistors were easily obtained with a Hewlett Packard model 4145A semiconductor parameter analyzer. Drain voltage was linearly ramped in 50 mV steps from 0 to 7 volts, applied gate voltages were 0, 1, 2,

and 3 volts, and the source and substrate were grounded. Figure 4-6 gives a top view of one set of wide channel transistors.

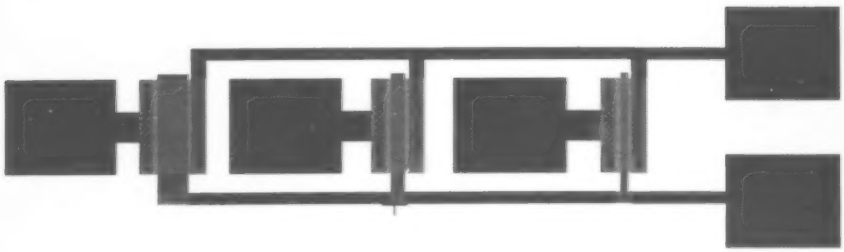


Figure 4-6. Top view of designed NMOS transistors.

4.2.2.2 Sub-threshold Leakage Current (I_{D0})

Ideally, the current in a NMOS transistor at sub-threshold gate voltages should be zero. In reality, a small I_D current flows at these voltages. This sub-threshold leakage current, I_{D0}, may be seen in a plot of I_D vs. V_{GS} for a transistor operating in the linear region. At low V_{GS} (below threshold), I_{D0} can be read directly from the I_D-V_{GS} curve.

4.2.2.3 Threshold Voltage (V_{T0})

To determine the threshold voltage of fabricated transistors, two methods were employed. The first was to bias the transistor into the saturation region of operation. By setting $V_{GS} = V_{DS}$ and $V_S = V_{BS} = GND$, a plot of $\sqrt{I_D}$ versus V_{GS} can be extrapolated to zero drain current to give the threshold voltage [40] as given by equation 1.11. A similar technique can be used to determine V_{T0} for transistors operating in the linear region. Here, a small constant voltage (50 mV) is placed on the drain, a linear ramping voltage on the gate, and the source and substrate are grounded. From a plot of I_D vs. V_{GS}, a straight line can be fitted to the point of maximum slope. This line is extrapolated back to $I_D = 0$ to give V_{T0}. From equation 1.10, it can be seen that $V_{T0} = V_{GS} - V_{DS}/2$.

4.2.2.4 Transconductance Parameter (k′)

The transconductance parameter, k′, from equation 1.11 can be obtained from a $\sqrt{I_D}$ -V_{GS} curve. This is done with the transistor in the saturation region. The slope of this curve is $\sqrt{\frac{k'W}{2L}}$ and k′ can be calculated by substituting in the values for W and L.

4.2.2.5 Channel Length Modulation Parameter (λ)

The channel length modulation parameter, discussed in section 1.1.2, can be measured by extrapolating the slope of I_D-V_{DS} curves in the saturation region of operation. From equation 1.12, at $I_D = 0$ the x-intercept is $1/\lambda$ V^{-1}.

5. Test Results and Discussion

For process development, wafers were processed in batches of 3, excluding test wafers (production batches would of course be larger). Two batches were carried to completion. While spot testing was carried out on most of the wafers, one wafer from each batch was selected for thorough characterization. The results from these wafers are discussed below.

Characterization data collected from both wafer and device level testing is presented. Wafer testing consisted of four point probe and diffusion junction depth measurements. Other measurements were taken on individual wafer die. To simplify the presentation, wafer die will be numbered from 1 to 16 and test results will be referenced by die number. Figure 5-1 shows the numbering scheme.

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Figure 5-1. Numbering scheme for wafer die.

Because of the large number of test structures and devices fabricated, only test results from specified die will be presented in this chapter.

5.1 Sheet Resistance and Junction Depth

The times and temperatures used for diffusion were estimated by assuming a final sheet resistance of 20 Ω/sq. and junction depth of 0.8 μm for the n-doped source and drains. These values were obtained for a similar process developed at Carleton University [29].

By definition, n-doping concentration at the junction depth is equal to the background doping concentration of the original p-type substrate. Therefore, equation 3.7 can be used to determine a suitable drive-in time and temperature. The following calculations were done by using equation 3.7 and Irvin curves [47]. Irvin published two types of curves relating dopant concentration and resistivity (or conductivity) for bulk silicon and diffused layers in silicon. Curves for diffused layers plot layer concentration versus average conductivity for a given initial background concentration. Curves for bulk silicon show the basic relationship between resistivity and impurity concentration which is non-linear since carrier mobility decreases at high concentrations.

Sheet resistance can be defined as $R_s = \overline{\rho}_s / x_j$ where $\overline{\rho}_s$ is average resistivity and x_j is the junction depth. For a desired R_s = 20 Ω/sq. with x_j = 0.8 μm, $\overline{\rho}_s$ = 1.60 x 10^{-3} Ω-cm. Using the curve for initial 1-10 Ω-cm p-type wafers, the impurity

concentration varies from about 1.60 x 10^{16} cm^{-3} for 1 Ω-cm wafers to 1.30 x 10^{15} cm^{-3} for 10 Ω-cm wafers. This will be called the background concentration, C_B. By using the Gaussian Irvin curve for $C_B = 10^{16}$ cm^{-3} and $\bar{\sigma}_s = 1/\bar{\rho}_s = 625\ (\Omega\text{-cm})^{-1}$, we find a surface concentration, C_{S2}, of about 2.25 x 10^{20} cm^{-3}. Solving equation 3.7 for D_2t_2, with $x = x_j$ and $C(x_j,t_2) = C_B$, gives

$$D_2t_2 = \frac{x_j^2}{4 \ln\left(C_{S2}/C_B\right)} \qquad (5.1)$$

Substitution of the above numbers gives D_2t_2 = 1.60 x 10^{-10} cm^2. To keep t_2 to a reasonable value, it was decided to perform drive-in at 1050°C. Data published by Ishikawa et al. [35] gives a value of about 1.50 x 10^{-13} cm^2/sec for D_2 at 1050°C. Solving for t_2 then gives a drive-in diffusion time of 1067 sec.

C_{S2} can now be used to determine the dose, Q, required to achieve the desired sheet resistance and junction depth. Since $Q = C_{S2}\sqrt{\pi D_2t_2}$, we have Q = 5.0 x 10^{15} cm^{-2}. Based on this calculated dose, data supplied by the manufacturer of the diffusion sources was then used to estimate an appropriate predeposition temperature, T_1, and time, t_1 [48]. Table 5-1 shows the calculation results for wafers of 1 and 10 Ω-cm resistivity.

Initial ρ_s (Ω-cm)	C_B (cm^{-3})	C_{S2} (cm^{-3})	t_2 (min-sec)	Q (cm^{-2})
1	1.6×10^{16}	2.3×10^{20}	17 min 47 sec	5.0×10^{15}
10	1.3×10^{15}	2.7×10^{20}	14 min 27 sec	5.4×10^{15}

Table 5-1. Initial estimates for diffusion times.

Experimentation with different predeposition and drive-in conditions produced the desired sheet resistance of 20 Ω/sq. with the times and temperatures mentioned in section 3.4.12. All subsequent impurity doping used these times and temperatures.

5.1.1 Four Point Probe

Measurements taken by four point probe after drive-in showed that the sheet resistance values varied greatly from those obtained with test runs under the same conditions. Table 5-2 summarizes the data collected on two diffusion test wafers which were placed in the same boat as full process wafers.

Wafer Position	Test Wafer #1 Sheet Resistance (Ω/sq.)	Test Wafer #2 Sheet Resistance (Ω/sq.)
1	45.2 n-type	40.4 n-type
2	61.0	44.2
3	65.9	63.9
4	68.3	54.6
5	62.3	54.1

Table 5-2. Four point probe measurements on diffusion test wafers.

Test wafer #1 was doped with a 4.0 ℓ/min flow of N_2 during predeposition and test wafer #2 used a 2.0 ℓ/min flow. To ensure that valid comparisons could be made between test and full process wafers, wafers with similar initial sheet resistances were doped together.

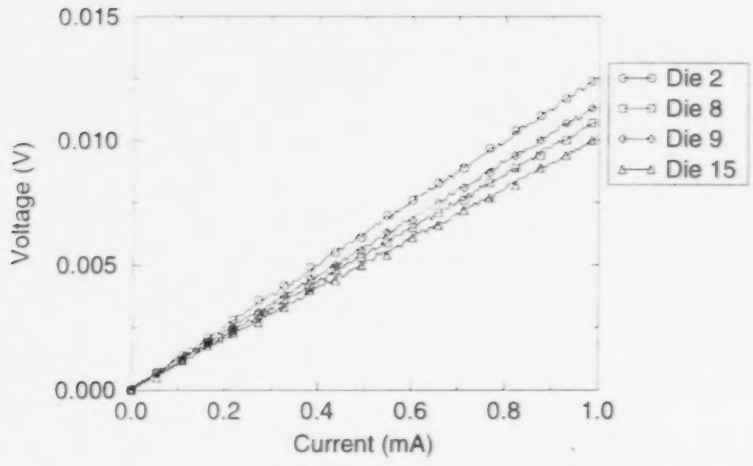
Results show that sheet resistance varied across the wafer. Lower sheet resistance in the middle of wafers indicate they were more heavily doped there. At first, it was hypothesized that too high of a nitrogen flow rate was used during predeposition diffusion. This would have caused turbulence at the wafer edges and affected doping there. However, a study performed by K. Harris at the AMC showed otherwise [49]. Experiments with the planar doping sources indicated that the sources themselves were at fault and not the gas flow rate used for predeposition.

Two main reasons for the variation in sheet resistance are possible. The first involves the poor quality of the PECVD deposited polysilicon described earlier. During predeposition, the high temperature caused portions of polysilicon film to blister off wafer surfaces. This may have contaminated the surface of adjacent dopant sources and affected the doping uniformity across wafers and across the diffusion boat. A second possibility lies in the frequency of use of the phosphorus doping sources. Wafer doping was not done on a regular basis and the storage time in between uses may have been sufficiently long to reduce the effectiveness of the

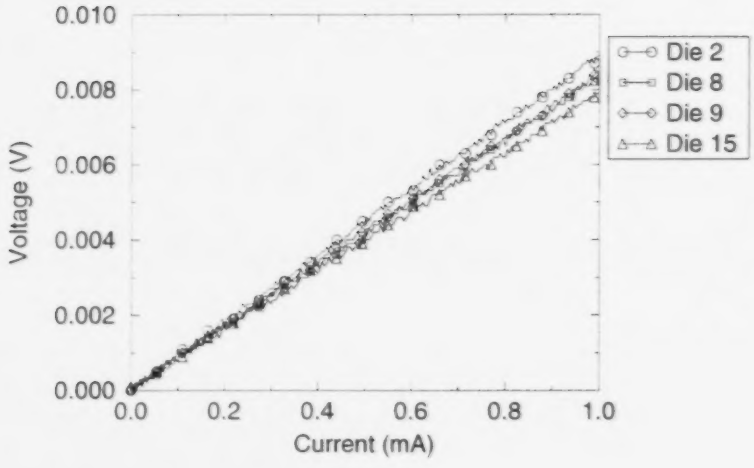
dopant sources. Although sources were reannealed according to manufacturer's specifications, this step may not have been adequately done. Initial trials using new dopant sources produced better results with improved uniformity.

5.1.2 Van der Pauw Measurements

Sheet resistance of fully processed wafers was determined by the van der Pauw technique. Measurements taken on two separately processed wafers are shown in figures 5-2 to 5-5. The graphs show the voltage difference V_1 - V_2 for current ranging from 0 to 1 mA for Greek and offset quadrate crosses. A straight line was fitted to the experimental data and sheet resistance was then calculated using equation 4.2. These values are summarized in tables 5-3 and 5-4.

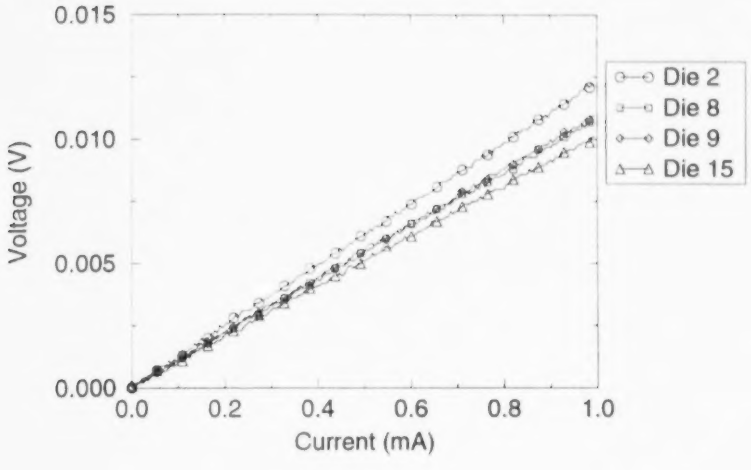


(a)

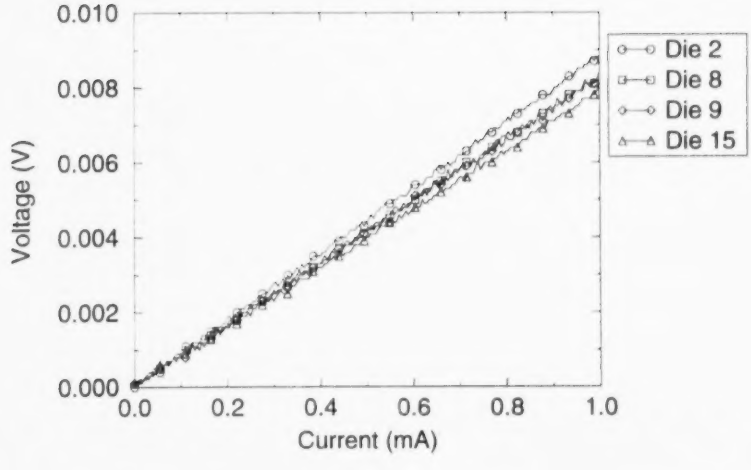


(b)

Figure 5-2. Diffused Greek crosses (a) wafer #1, (b) wafer #2.

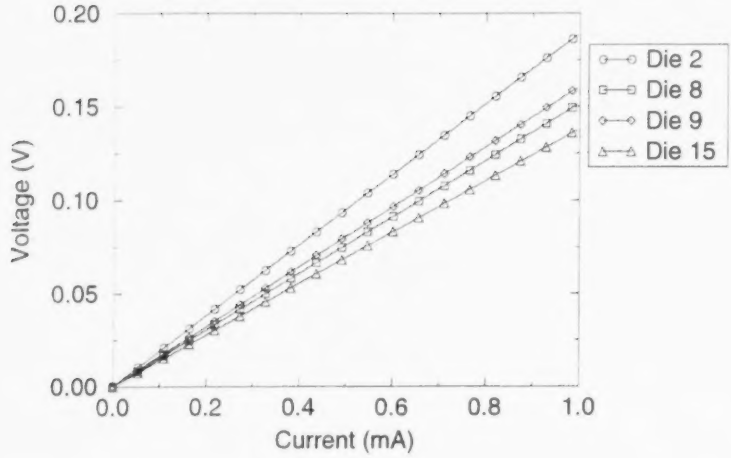


(a)

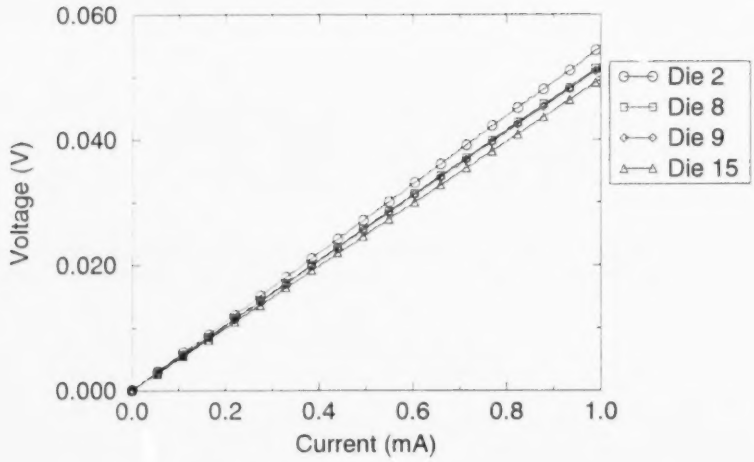


(b)

Figure 5-3. Diffused offset quadrate crosses (a) wafer #1, (b) wafer #2.

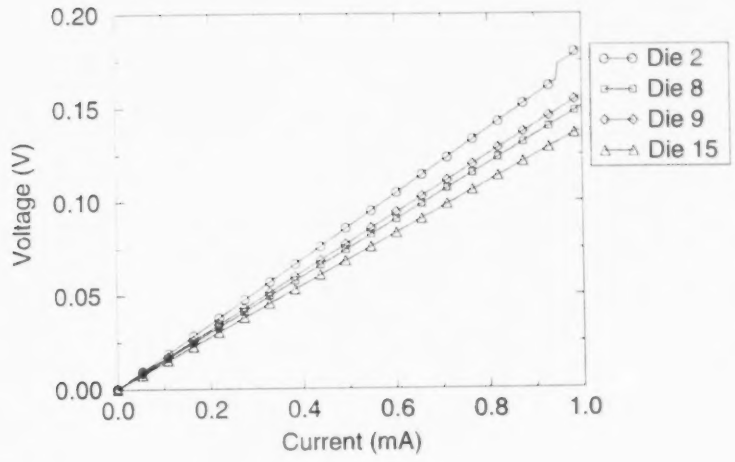


(a)

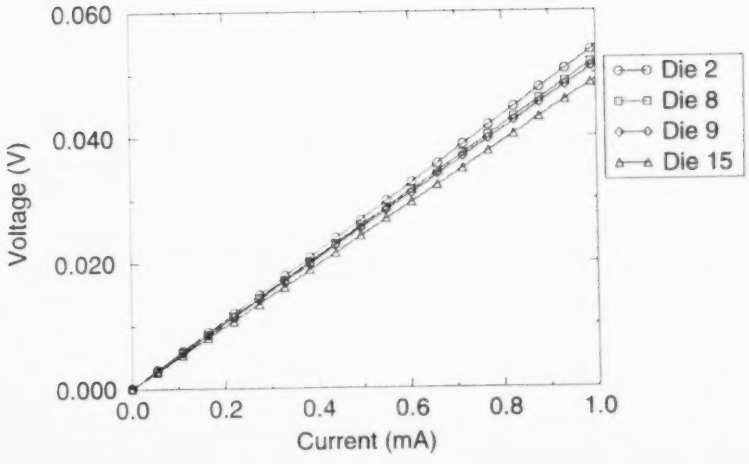


(b)

Figure 5-4. Doped polysilicon Greek crosses (a) wafer #1, (b) wafer 2.



(a)



(b)

Figure 5-5. Doped polysilicon offset quadrate crosses (a) wafer #1, (b) wafer #2.

Die	Wafer #1 Greek cross	Wafer #2 Greek cross	Wafer #1 Offset cross	Wafer #2 Offset cross
2	56.6 Ω/sq.	40.2 Ω/sq.	55.5 Ω/sq.	39.8 Ω/sq.
8	48.8	37.4	48.8	37.4
9	51.5	37.9	49.6	37.1
15	45.5	35.6	45.8	35.4

Table 5-3. Van der Pauw sheet resistance of diffused crosses.

Die	Wafer #1 Greek cross	Wafer #2 Greek cross	Wafer #1 Offset cross	Wafer #2 Offset cross
2	856.2 Ω/sq.	247.7 Ω/sq.	787.7 Ω/sq.	245.8 Ω/sq.
8	686.4	234.6	681.6	236.6
9	727.5	233.1	708.8	233.3
15	625.5	224.7	627.1	222.1

Table 5-4. Van der Pauw sheet resistance of doped polysilicon crosses.

The table values show lower sheet resistance at wafer centres. This is consistent with four point probe measurements taken on the test diffusion wafers. There also seems to be a variation in sheet resistance between process wafers. This can be attributed to the unpredictable nature of the dopant sources. Polysilicon structures exhibited a large sheet resistance. The reasons behind this may lie in the inferior quality of the PECVD deposited polysilicon and contamination of the dopant sources by polysilicon.

5.1.3 Diffusion Junction Depth

Diffusion junction depth was determined using the method described in section 4.1.5. A scanning electron microscope photograph, shown in figure 5-6, reveals a junction depth of about 1 μm. The resolution of this method is not very good. If greater resolution were required, junction depths could be magnified by beveling the sample before staining or a method such as SIMS (secondary ion mass spectrometry) could be used.

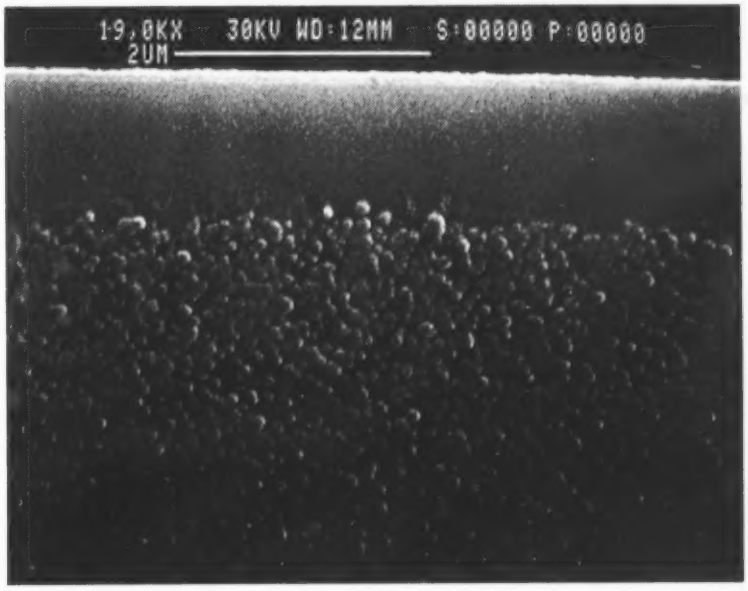


Figure 5-6. SEM photograph of junction depth on a diffused wafer.

5.1.4 Comparison against Diffusion Theory

From the data presented, it can been seen that there is a general trend of increasing sheet resistance radially outwards from wafer centres. Heavier doping at wafer centres may be verified from the data with calculations for dopant dose and surface concentration. By using equations 3.6 and 3.7 and Irvin curves, values for Q and C_{S2} may be calculated. By multiplying the measured junction depth of 1 μm and the average sheet resistance of van der Pauw crosses, an average resistivity, and hence, conductivity, may be determined. From Irvin curves, the surface concentration at that conductivity can be found. From this value of C_{S2}, equations 3.6 and 3.7 are used to determine the value for Q. To see if the Gaussian drive-in diffusion model accurately predicts the measured junction depth, the dose and substrate background concentration can be substituted into equation 3.6 to calculate a junction depth. Tables 5-5 and 5-6 summarize calculations based on diffusion measurements.

	Wafer # 1, ($C_B \sim 7.0 \times 10^{15}$ cm^{-3})		
Die	Q (cm^{-2})	C_{S2} (cm^{-3})	x_j (μm)
2	1.16×10^{15}	4.00×10^{19}	0.967
8	1.38×10^{15}	4.75×10^{19}	0.976
9	1.31×10^{15}	4.50×10^{19}	0.973
15	1.60×10^{15}	5.50×10^{19}	0.984

Table 5-5. Diffusion calculations for wafer #1.

Wafer # 2, (C_B ~ 1.7 x 10^{15} cm^{-3})			
Die	Q (cm^{-2})	C_{S2} (cm^{-3})	x_j (μm)
2	2.04 x 10^{15}	7.00 x 10^{19}	1.07
8	2.18 x 10^{15}	7.50 x 10^{19}	1.07
9	2.18 x 10^{15}	7.50 x 10^{19}	1.07
15	2.33 x 10^{15}	8.00 x 10^{19}	1.08

Table 5-6. Diffusion calculations for wafer #2.

The calculations show very good agreement with the measured junction depth. It can also been seen that the doses from different parts of the wafer correspond to the measured sheet resistances. Die which are closest to the wafer centres show higher doping concentrations.

5.2 Contact Resistance

5.2.1 Contact Strings

Data collected from contact string measurements are shown in figures 5-7 and 5-8 for process wafers #1 and #2. The curves exhibit non-linear behaviour at higher voltages, so to determine the slope a straight line was fitted to points at voltages less than 1.5 V. Using the inverse slope, individual contact resistance was determined by subtracting the total resistance of the twenty doped resistors and dividing by the

number of contacts. Resistor values were calculated from an average sheet resistance between Greek and offset crosses, as determined in section 5.1.2. The values are given in tables 5-7 and 5-8.

Die	Diffused String		Doped Polysilicon String	
	Total resistance (kΩ)	Ind. contact resistance (Ω)	Total resistance (kΩ)	Ind. contact resistance (kΩ)
2	6.73	91.8	597	13.8
8	5.60	73.3	435	9.94
9	5.51	68.7	430	9.77
15	5.17	66.9	337	7.57

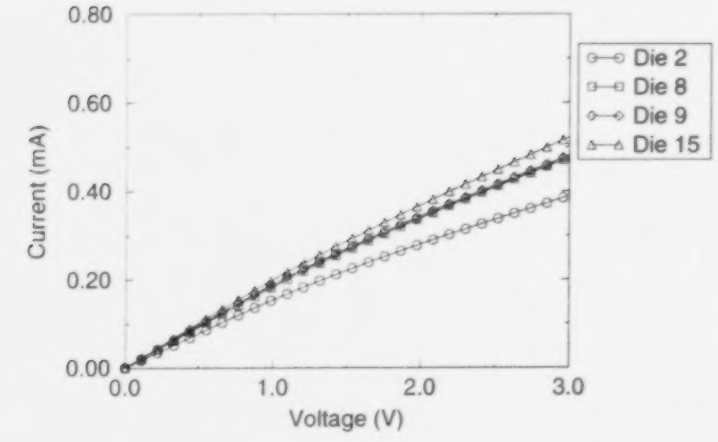
Table 5-7. Contact resistance of wafer #1 using contact strings.

Die	Diffused String		Doped Polysilicon String	
	Total resistance (kΩ)	Ind. contact resistance (Ω)	Total resistance (kΩ)	Ind. contact resistance (Ω)
2	8.77	165	47.0	838
8	6.81	119	35.9	576
9	7.09	126	37.6	621
15	5.99	101	30.6	460

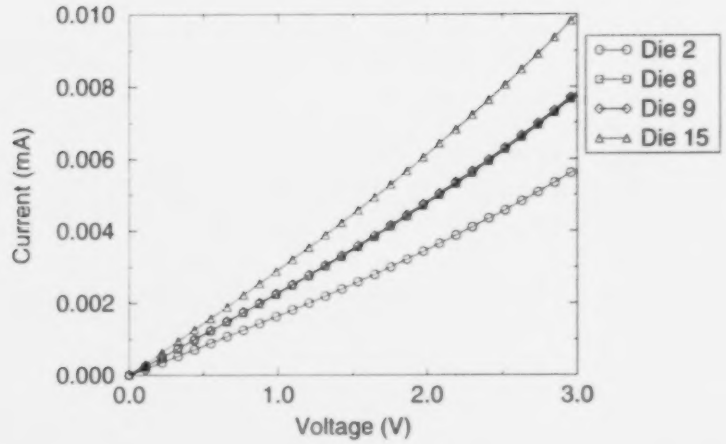
Table 5-8. Contact resistance of wafer #2 using contact strings.

Because of the non-linear behaviour of the data, values obtained by the partial curve fits must be considered questionable. The contact string method is considered only a rough estimate to contact resistance [40]. However, it may be seen that the more heavily doped regions of the wafer gave lower contact resistance. Contacts between aluminum and n-doped silicon are inherently non-ohmic. By doping to a

sufficiently high concentration (~ 5 x 10^{19} cm^{-3}), ohmic behaviour can be achieved [50,51]. Since the polysilicon was more lightly doped, it gave higher contact resistance values. Values for contact resistance were somewhat higher than expected but were within an order of magnitude of published values [50]. Possible reasons for this discrepancy include contamination due to inadequate cleaning before dielectric deposition or an unoptimized contact anneal. There also exists the difficulty of trying to characterize nonlinear behaviour with a linear model.

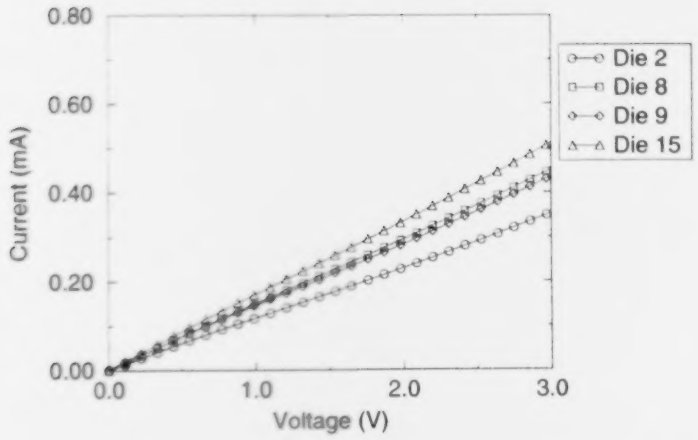


(a)

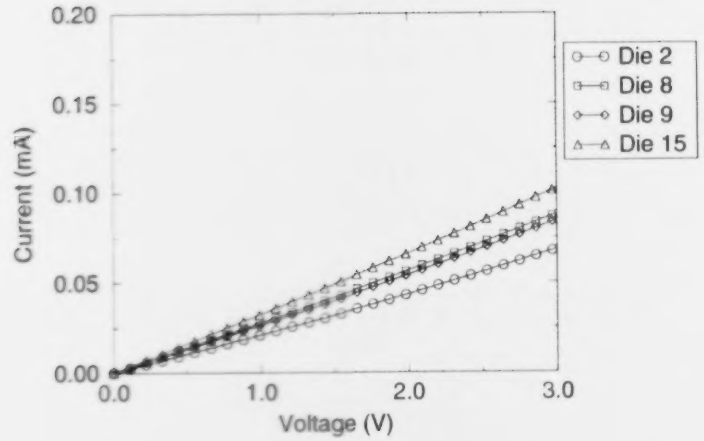


(b)

Figure 5-7. Contact string measurements on wafer #1 (a) diffused string, (b) doped polysilicon string.



(a)



(b)

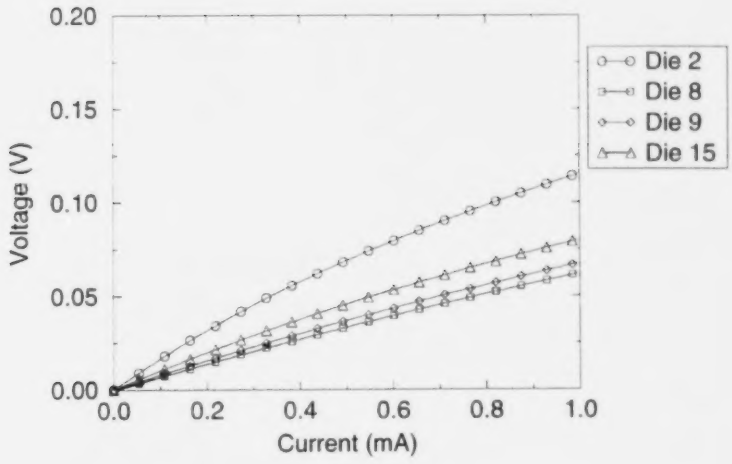
Figure 5-8. Contact string measurements on wafer #2 (a) diffused string, (b) doped polysilicon string.

5.2.2 Four Terminal Kelvin Test Structures

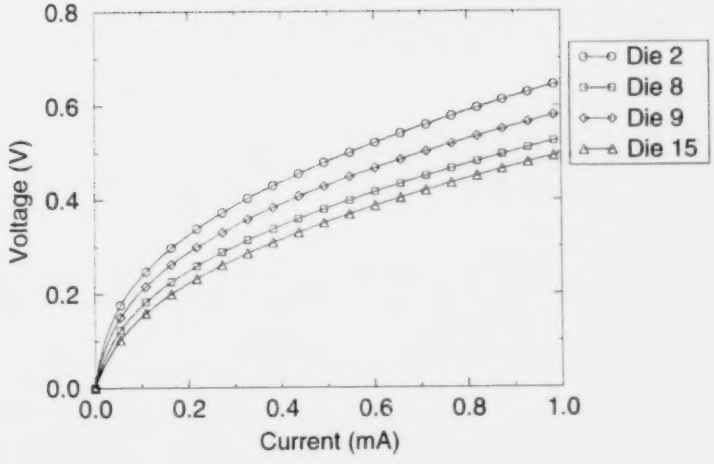
Data in figures 5-9 and 5-10 plot the voltage difference V_3 - V_4 against current for measurements taken from Kelvin test structures on two wafers. It can be seen that at low currents contact behaviour is non-ohmic. To determine contact resistance from these curves, a straight line was fitted to points at a current greater than 5 mA. This part of the curve can be seen to be more linear. Results obtained from the graphs are summarized in table 5-9.

Die	Wafer #1 Contact Resistance		Wafer #2 Contact Resistance	
	Diffused (Ω)	Doped Polysi (Ω)	Diffused (Ω)	Doped Polysi (Ω)
2	92.7	331	56.2	153
8	57.7	290	43.0	131
9	61.4	302	48.7	140
15	69.1	285	39.6	117

Table 5-9. Calculated contact resistance using Kelvin test structures.

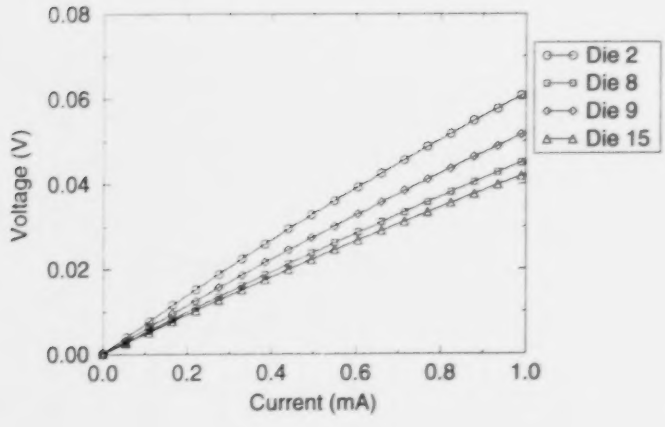


(a)

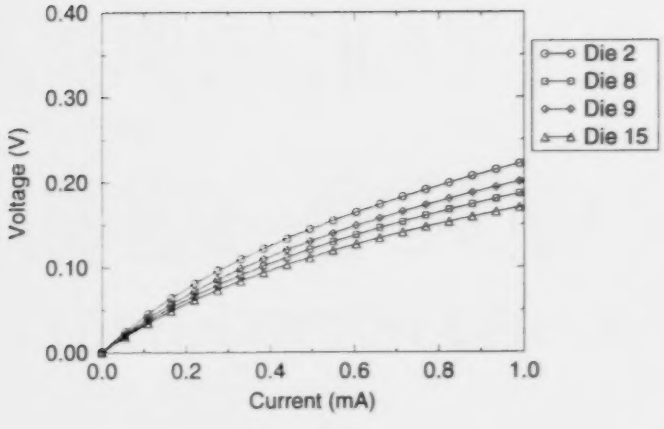


(b)

Figure 5-9. Kelvin contact resistance on wafer #1 (a) diffused, (b) doped polysilicon.



(a)



(b)

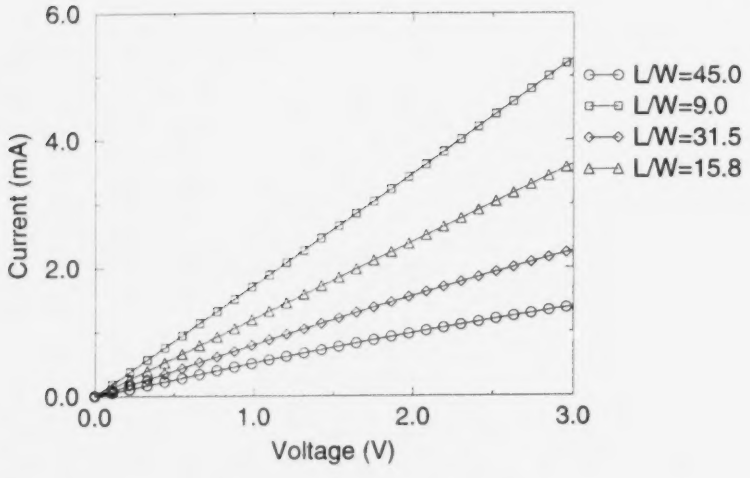
Figure 5-10. Kelvin contact resistance on wafer #2 (a) diffused, (b) doped polysilicon.

Contact resistance values from Kelvin measurements again show a dependence on doping concentration. Curves from wafer #2 show a slightly more linear characteristic than those from wafer #1 due to higher doping concentration. Again, the polysilicon on both wafers was insufficiently doped to yield ohmic behaviour. However, for the 7.5 μm square contacts to diffused silicon, values for contact resistance were only a few tens of ohms higher than those expected [50].

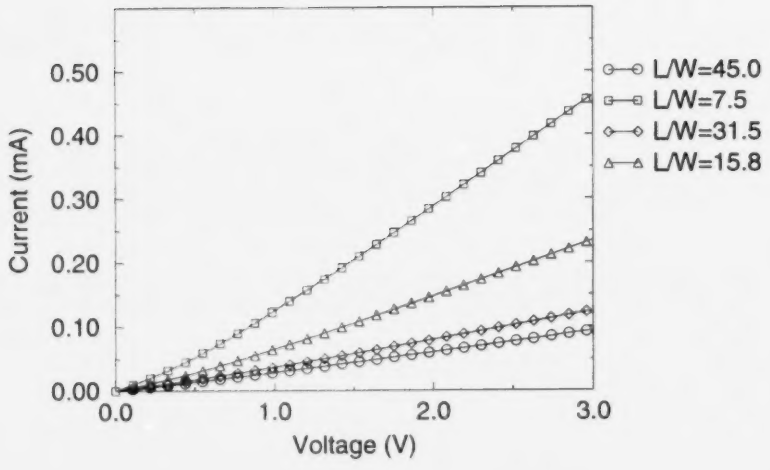
The discrepancy in value may be due to the Kelvin structure itself. Contacts which are not the exact same size as the underlying diffused region allow current to flow around the contact. This has been found to contribute an additional resistance to the measured contact resistance [52].

5.3 Resistors

Two sets of resistor curves are shown in figures 5-11 and 5-12. The figures show current-voltage characteristics for 8 diffused and doped polysilicon resistors. Resistor values were obtained from straight line curve fits to the data. These are given in table 5-10. Expected resistance values were calculated using an average sheet resistance calculated from the van der Pauw crosses of section 5.1.2.

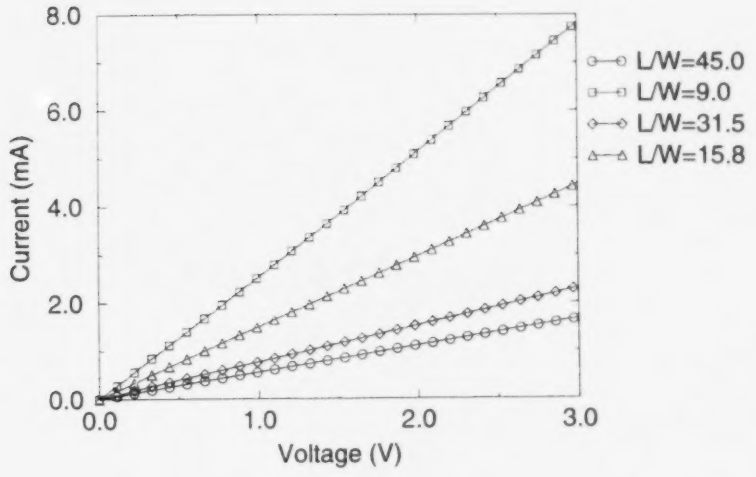


(a)

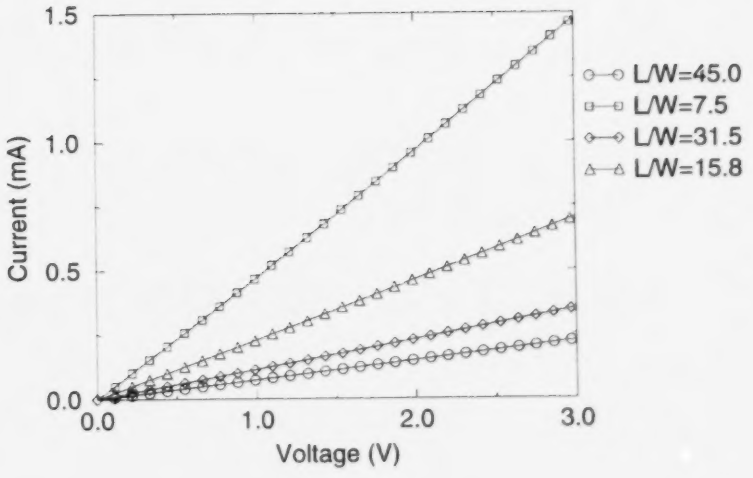


(b)

Figure 5-11. Wafer #1 resistors (a) diffused, (b) doped polysilicon.



(a)



(b)

Figure 5-12. Wafer #2 resistors (a) diffused, (b) doped polysilicon.

Resistor type and size	Wafer #1 Die 9 Resistance (kΩ)		Wafer #2 Die 8 Resistance (kΩ)	
	Expected	Measured	Expected	Measured
Poly, L/W=45	32.3	30.4	10.6	13.3
Poly, L/W=7.5	5.39	5.84	1.77	1.97
Diff, L/W=45	2.27	2.33	1.68	1.80
Diff, L/W=9	0.455	0.561	0.337	0.376
Poly, L/W=31.5	22.6	22.6	7.42	8.52
Poly, L/W=15.8	11.3	11.6	3.72	4.19
Diff, L/W=31.5	1.59	1.38	1.18	1.30
Diff, L/W=15.8	0.799	0.827	0.591	0.671

Table 5-10. Resistor values for fully processed wafers #1 and #2.

Measured and predicted resistor values varied slightly from one another but demonstrated the ability to reliably produce resistors in the 300 Ω - 30 kΩ range. Resistance measurements were basically done by a two point probe method, so this would introduce the effects of contact resistance into the measurement. Different contact sizes were used in the resistor design, so the exact value of contact resistance could not be determined. Further errors may come from using an average sheet resistance to calculate expected resistance values.

5.4 Transistors

The primary goal, of course, of an NMOS process is to produce usable transistors. Therefore, considerable effort was expended to characterize and understand transistors produced. Comparisons were made to both analytic and SPICE models.

To minimize the errors due to lateral diffusion, transistor data shown in this section was taken from transistors with the largest designed channel widths and lengths.

5.4.1 I_D - V_{DS} Curves

The following figures show the dc characteristics from six transistors tested on two separate batches of wafers. All transistors have a gate width of 150 μm and vary only in length. Analysis of these curves to determine individual transistor parameters is given in subsequent sections.

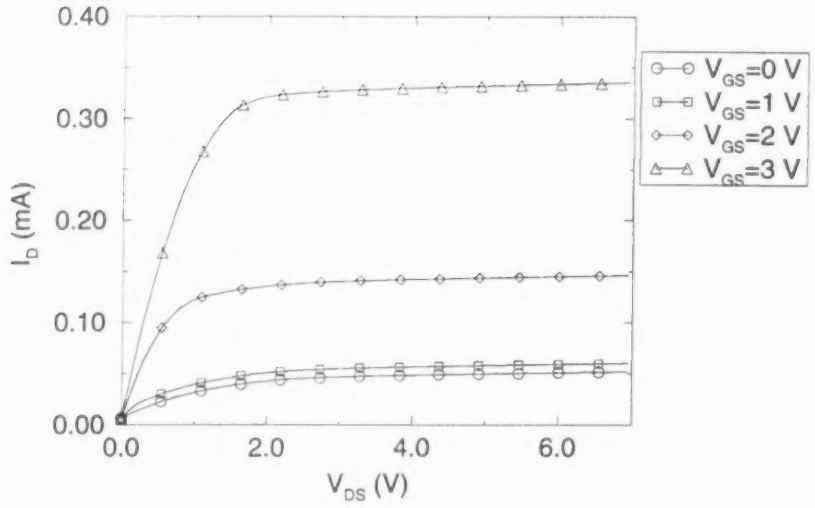


Figure 5-13. I-V characteristics for a transistor with L = 40 μm (wafer #1).

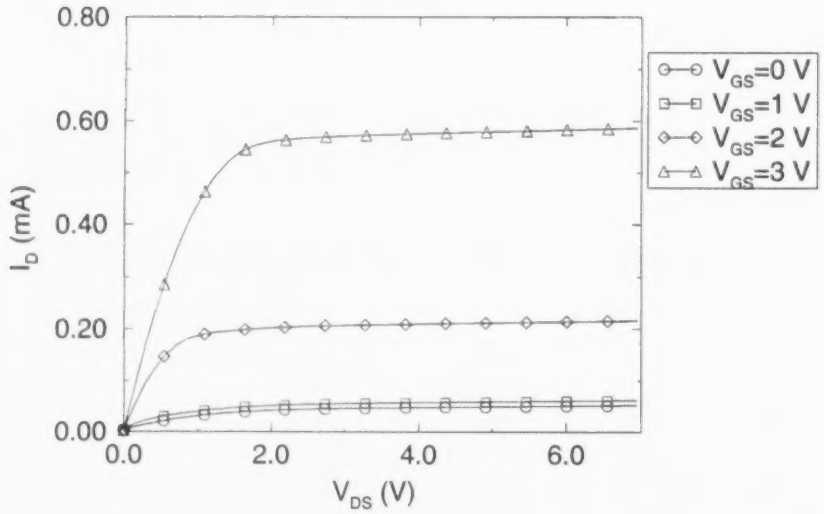


Figure 5-14. I-V characteristics for a transistor with L = 20 μm (wafer #1).

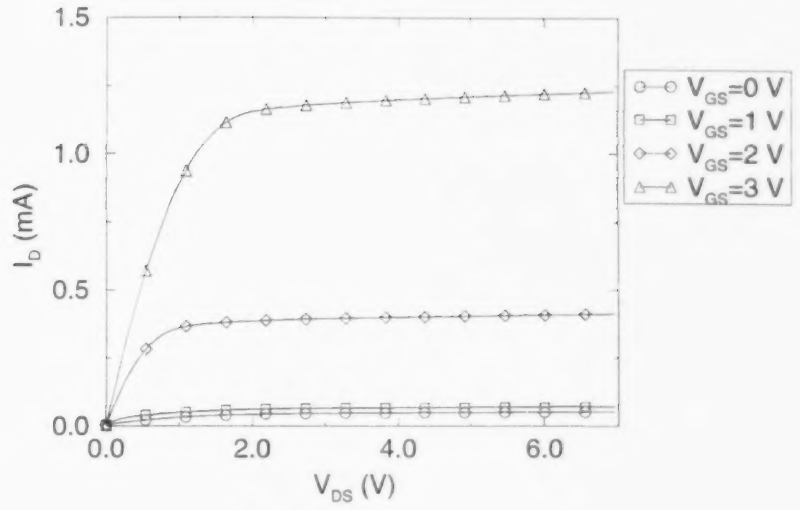


Figure 5-15. I-V characteristics for a transistor with L = 10 µm (wafer #1).

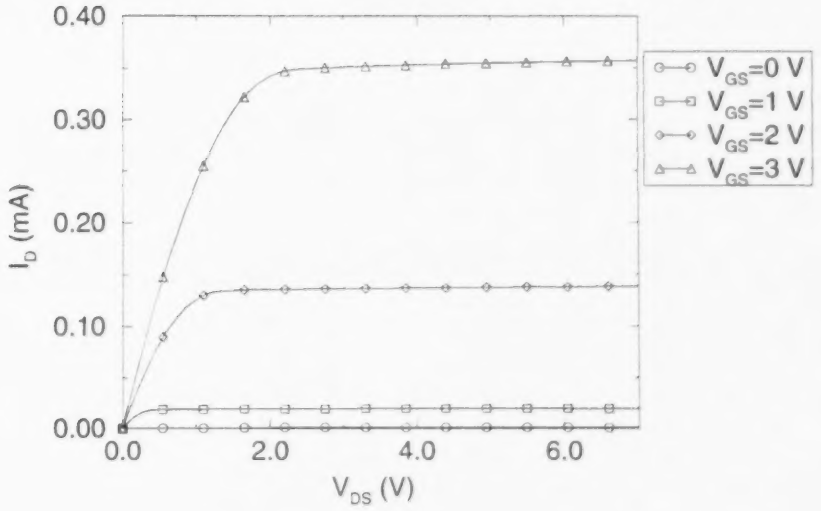


Figure 5-16. I-V characteristics for a transistor with L = 40 µm (wafer #2).

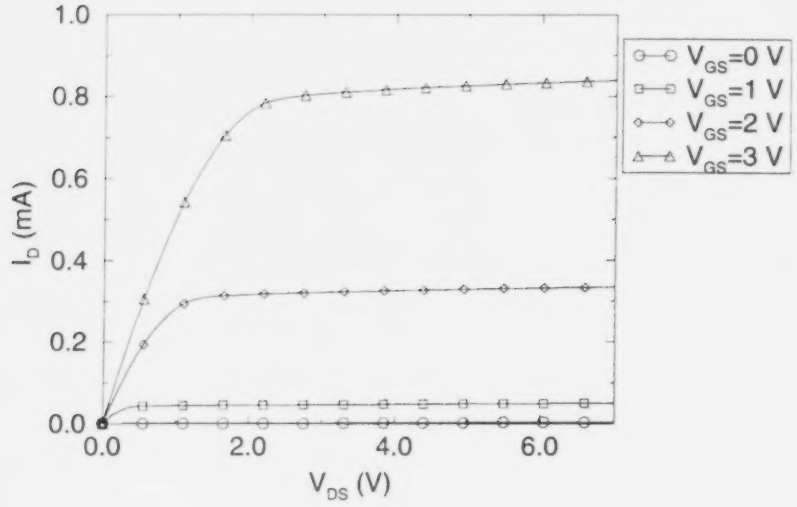


Figure 5-17. I-V characteristics for a transistor with L = 20 μm (wafer #2).

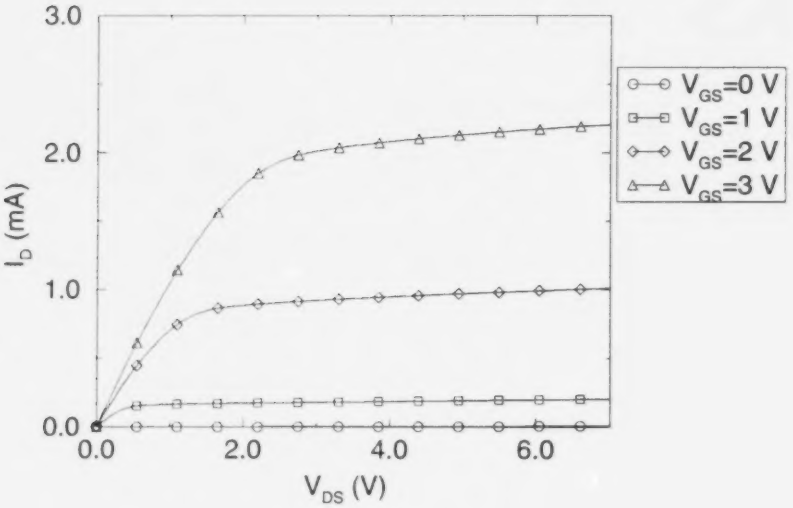


Figure 5-18. I-V characteristics for a transistor with L = 10 μm (wafer #2).

5.4.2 Sub-threshold Leakage Current (I_{D0})

Leakage currents were determined from transistors operating in the linear mode of operation. These curves are shown in the following section. On the first process wafer, leakage currents of 150 micron wide transistors were found to be 8.1, 6.2, and 7.5 μA for transistors 40, 20, and 10 microns in length. Leakage currents of 0.54, 0.76, and 0.41 μA were found on similarly sized transistors on the second process wafer to be tested. The leakage levels for wafer #1 are quite substantial and may be due to wafer contamination. Leakage currents for wafer #2 seem to be acceptable.

5.4.3 Threshold Voltage (V_{T0})

Threshold voltage measurements were taken from transistors operating in the saturation and linear regions. Figures 5-19 to 5-22 show I_D - V_{GS} curves for the six transistors tested in section 5.4.1. Threshold voltage was extrapolated from the x-intercept of straight line curve fits done on the linear part of the curves.

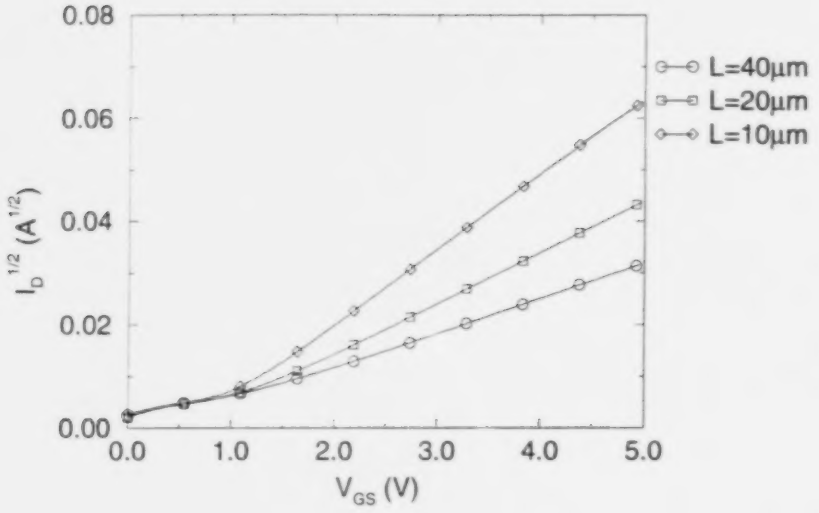


Figure 5-19. Saturation region transistor operation (W = 150 µm) on wafer #1.

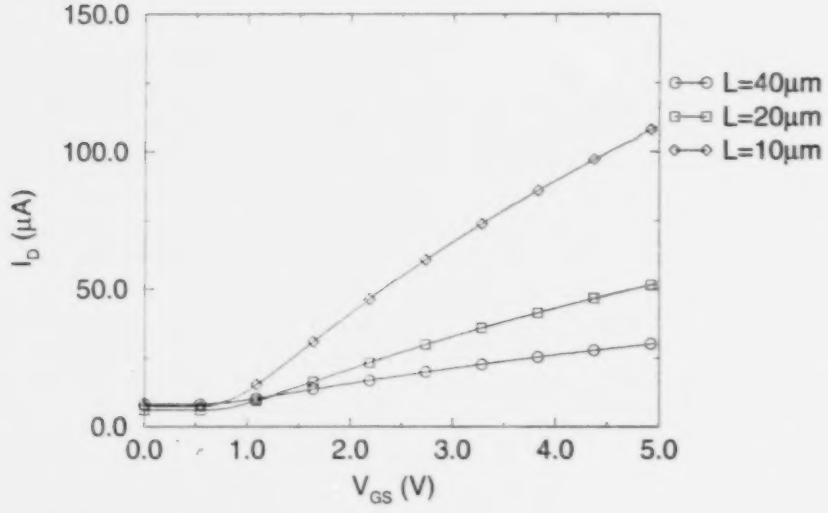


Figure 5-20. Linear region transistor operation (W = 150 µm) on wafer #1.

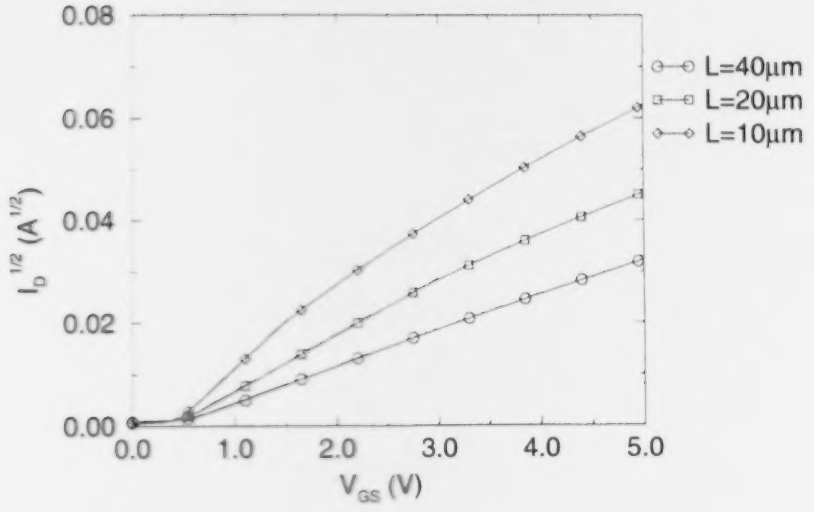


Figure 5-21. Saturation region transistor operation (W = 150 µm) on wafer #2.

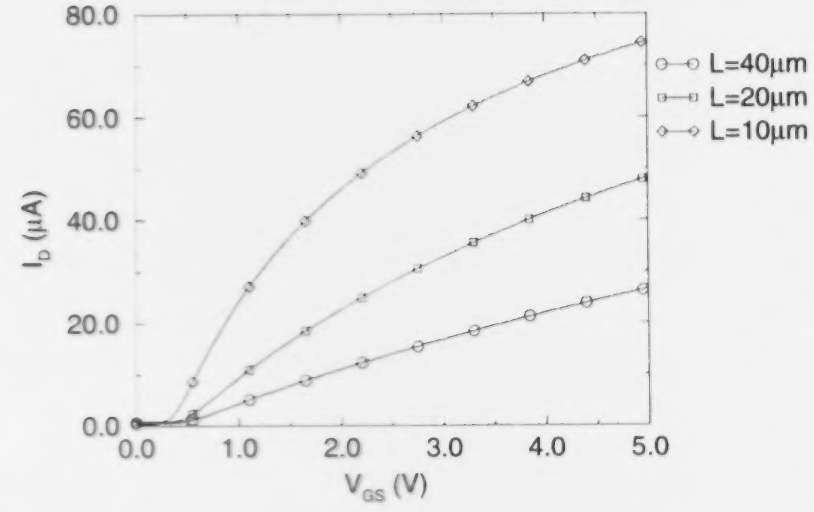


Figure 5-22. Linear region transistor operation (W = 150 µm) on wafer #2.

Threshold voltage determined from the above graphs is tabulated below.

Transistor Length	Wafer #1 Threshold Voltage (V) Saturation Region	Linear Region	Wafer #2 Threshold Voltage (V) Saturation Region	Linear Region
40 μm	0.268	0.556	0.371	0.379
20 μm	0.561	0.359	0.378	0.389
10 μm	0.636	0.557	0.386	0.300

Table 5-11. Extrapolated threshold voltage for transistors 150 μm wide.

Extrapolated threshold voltages indicate that fabricated transistors were enhancement mode transistors ($V_T > 0$ V). Values shown in the table above were obtained from "best fit" curves to the data. This is not an exact technique and is subject to interpretation. Even so, it can be seen that in the case of wafer #2, fairly consistent values can be obtained for both the linear and saturation modes. Variations in wafer #1 are likely due to complications due to the high leakage currents in that wafer. Both methods of extrapolating threshold voltage are sensitive to series resistance and mobility degradation [40]. Since these values were not determined in this work, it cannot be said for certain what effect they had on the measurements.

Due to the uncertainty in several of the terms (especially Q_{ox}) in equation 1.3, it is difficult to predict a value for V_T from process conditions. The observed value of about 0.37 V is a bit below that used in commercial devices which rely on an extra

threshold adjust implant. It is, however, likely suitable for MOS circuits without requiring a substrate bias to increase V_T according to equation 1.4.

5.4.4 Transconductance Parameter (k´)

Slopes of the $\sqrt{I_D}$ - V_{GS} curves presented earlier were used to determine values for k´. The results are summarized in table 5-12 below.

Transistor Size (W = 150 µm)	Wafer #1 k' (µA/V^2)	Wafer #2 k' (µA/V^2)
L = 40 µm	24.0	27.2
L = 20 µm	26.1	31.7
L = 10 µm	28.5	44.5

Table 5-12. Experimental values of k'.

For the gate oxide thicknesses (500 Å for wafer #1, 800 Å for wafer #2) used here, the expected values of k´ would be 35 µA/V^2 for wafer #1 and 22 µA/V^2 for wafer #2. This was assuming $\overline{\mu}_n = 500\,cm^2/V\cdot s$ and $\varepsilon_{ox} = 3.9\varepsilon_o$. The differences between estimated and observed values are likely due to variations in thickness and permittivity of the gate oxide.

5.4.5 Channel Length Modulation Parameter (λ)

The channel length modulation parameter was calculated from experimental data by fitting a straight line to the V_{GS} = 3 V curve in the I-V curves presented in section 5.4.1. For transistors with a width of 150 µm and lengths of 40 µm, 20 µm, and 10 µm, the modulation parameter was 5.40 x 10^{-3} V^{-1}, 6.35 x 10^{-3} V^{-1}, and 9.11 x 10^{-3} V^{-1}, respectively, for the first full process wafer. The second process wafer produced values of 3.67 x 10^{-3} V^{-1}, 9.06 x 10^{-3} V^{-1}, and 21.0 x 10^{-3} V^{-1} for the same sized transistors.

These values are consistent with reported values of 5 - 20 x 10^{-3} V^{-1} [53]. The increase with shorter channel length is to be expected as channel shortening due to depletion becomes more important.

5.4.6 Comparison of I_D - V_{DS} Curves with SPICE

The electrical parameters extrapolated from the data in this section were entered into SPICE simulations to generate simulated I_D - V_{DS} curves. Values for threshold voltage (V_{T0}), modulation parameter (λ), transconductance parameter (k'), and electrostatic potential (ϕ_F) were required in the simulations. The curves in figures

5-23 to 5-28 compare the resulting level 1 SPICE plots against measured curves for gate voltages of 0, 1, 2, and 3 volts. SPICE curves are shown as dashed lines.

Measured dc transistor characteristics compare well with the SPICE plots. Large differences are seen in the low gate voltage (weak inversion) curves for wafer #1 due to the large zero gate voltage leakage drain current. This current can also be seen to increase the drain current at V_{GS} = 1 V, where the drain current is a combination of real drain current and the leakage current. Because the theory presented in this work does not accurately model effects such as this, the origin of this current can only be hypothesized and not proven. It is quite possible that there is a positive charge in the gate oxide which is inducing a channel to exist even at zero gate voltage. However, this would imply that these devices should be depletion mode devices. Linear region curves show that this is not the case. Another possibility to this leakage current may be an unusually large reverse-junction leakage from the drain to substrate. These hypotheses were not verified and may be the subject for further research.

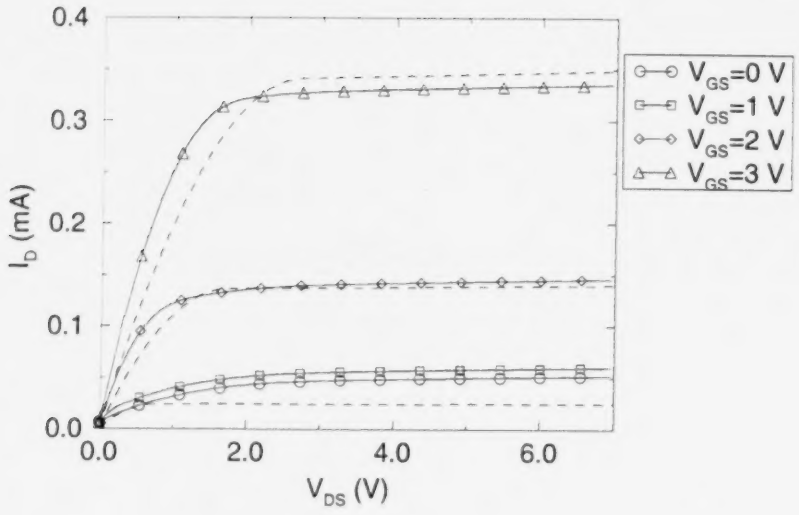


Figure 5-23. SPICE I_D - V_{DS} curves for a transistor with L = 40 μm (wafer #1).

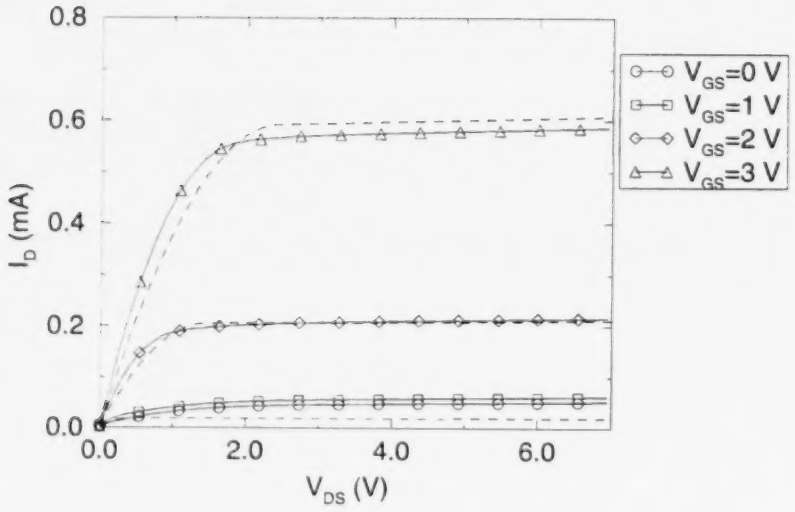


Figure 5-24. SPICE I_D - V_{DS} curves for a transistor with L = 20 μm (wafer #1).

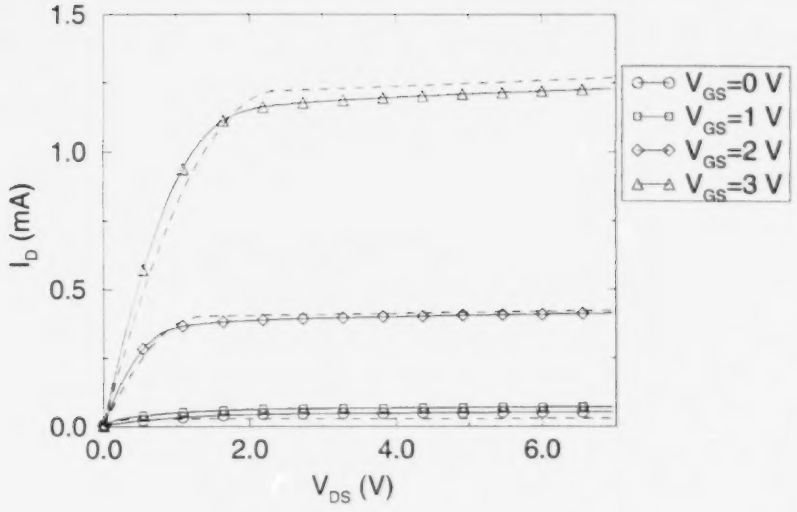


Figure 5-25. SPICE I_D - V_{DS} curves for a transistor with L = 10 µm (wafer #1).

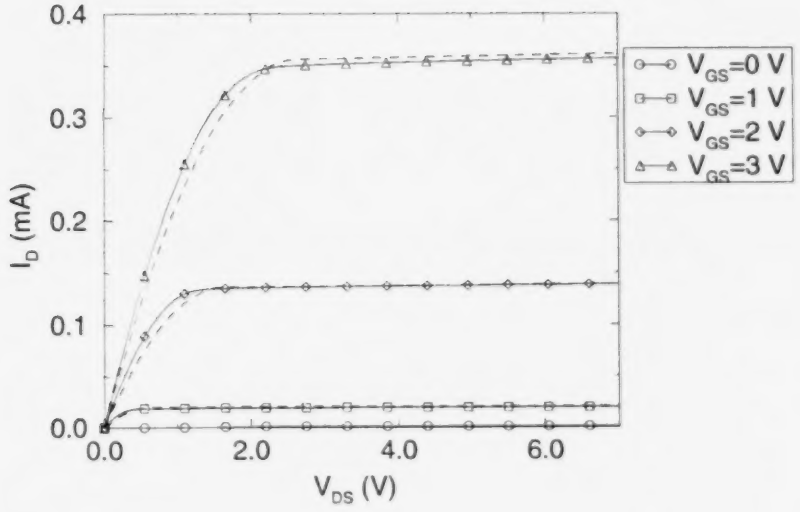


Figure 5-26. SPICE I_D - V_{DS} curves for a transistor with L = 40 µm (wafer #2).

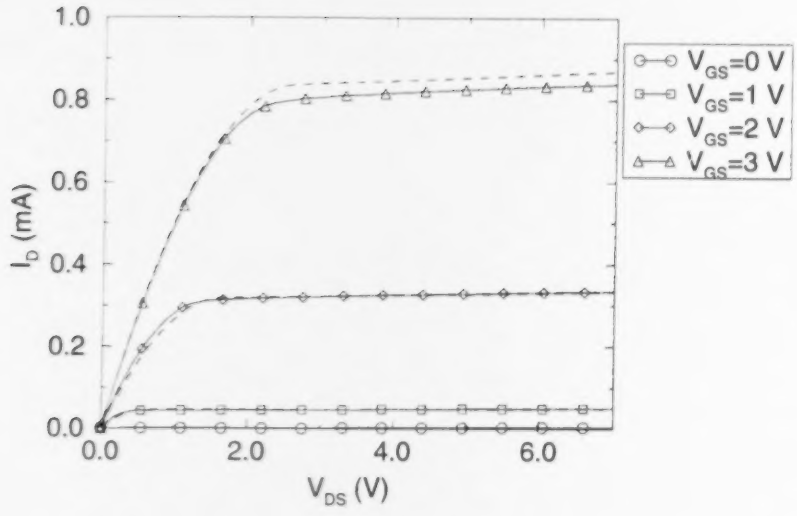


Figure 5-27. SPICE I_D - V_{DS} curves for a transistor with L = 20 μm (wafer #2).

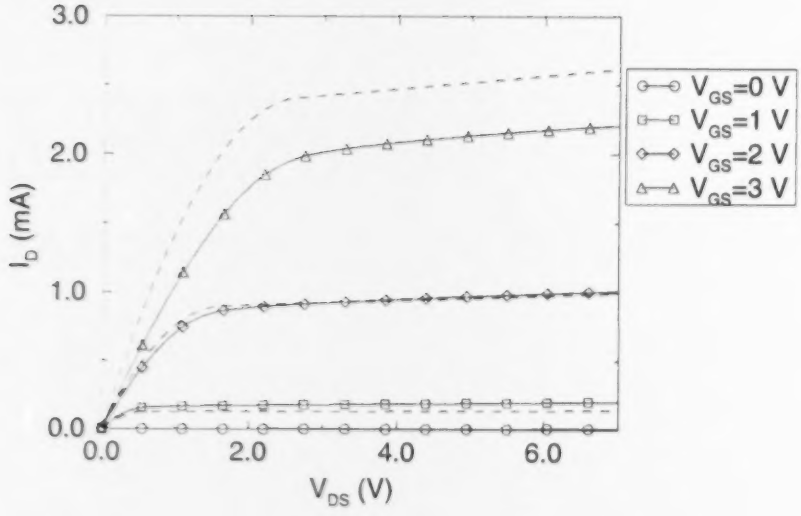


Figure 5-28. SPICE I_D - V_{DS} curves for a transistor with L = 10 μm (wafer #2).

One other difference between measured devices and SPICE simulations occurs at V_{GS} = 3 V. At this gate voltage, it may be seen that the drain currents on fabricated devices are slightly smaller than those predicted by SPICE. This effect can be explained by an overestimation of the transconductance, k´, or by mobility degradation [54]. Mobility degradation occurs at higher gate voltages when the larger normal electric field across the gate "pushes" electrons more towards the surface and reduces the effective mobility. Hence, I_D is reduced for these gate voltages.

An overestimation of the k´ parameter is clearly evident in figure 5-28. The effect is seen by a large difference between the simulated and measured curve for V_{GS} = 3 V. A value of 44.5 $\mu A/V^2$ was used for simulations of this transistor. As seen in table 5-12, this value was significantly higher than those obtained for other transistors on the same die. SPICE simulations with lower values of k´ produced better results.

6. Conclusions

The main goal of this work was to develop a process for fabricating NMOS transistors. Using the facilities available at the Alberta Microelectronic Centre, this goal was successfully accomplished. A four mask process utilizing PECVD deposited silicon nitride for LOCOS isolation, a self aligned PECVD deposited polysilicon gate, dry oxidation gate dielectric, solid source diffusion doping, and single level metallization was successfully implemented.

Process development included determining satisfactory conditions for oxidation, lithography, etching, and semiconductor doping. Film deposition parameters were already determined by previous experience and depositions were done based on these "process recipes". Early estimates to other process conditions were verified by experimentation.

Test structures designed to monitor process conditions included van der Pauw structures to measure sheet resistance and Kelvin test structures and contact strings to measure contact resistance. Sheet resistance and contact resistance were found to vary across wafer surfaces and between process runs. Doping non-uniformity was discovered to be the result of suspect dopant sources. New dopant sources appeared to solve this problem.

Fabricated devices were not as highly doped as desired. This was verified with sheet resistance and junction depth measurements. By assuming a Gaussian concentration profile for drive-in diffusion, a value from 4 x 10^{19} - 8 x 10^{19} cm^{-3} was calculated for surface dopant concentration. Doses were found to be about 1 - 2 x 10^{15} cm^{-2} and junction depths were about 1 μm.

Insufficient doping resulted in non-ohmic behaviour in metal-semiconductor contacts, particularly in doped polysilicon. However, there was less than an order of magnitude difference between measured values and published values for the 7.5 μm square contacts examined in this work.

Despite these problems, it was shown that enhancement mode transistors with good electrical characteristics could be fabricated. Transistors of varying size were successfully fabricated and tested. Wafers from the first batch of fabricated transistors exhibited large leakage currents. Nevertheless, usable devices with threshold voltages of 0.4 V were obtained without requiring a substrate voltage adjustment. Transconductance parameters from 24 - 40 $\mu A/V^2$ were observed, channel length modulation parameter values were in the 10^{-3} V^{-1} range, and leakage currents were generally found to be about 267 $\mu A/mm^2$ of transistor drain area. Comparisons of transistor dc characteristics with SPICE models yielded excellent agreement.

In addition to transistors, diffused and polysilicon resistors with values from 300 Ω - 30 kΩ were successfully fabricated. Other resistor values could have been obtained by varying the doping levels or changing the resistor dimensions.

Devices show promise for use in analog or digital circuits. Low cost and in-house control make it possible to integrate these devices with micromachined devices already produced at the AMC.

Future work may be based on optimizing process conditions. This could include semiconductor doping and PECVD deposited polysilicon. Problems with contact resistance could be solved by increasing the doping levels. As well, improved lithography and etching techniques could be used to yield smaller devices. Smaller devices would be faster but might require shallower junction depths. If so, an alternative to diffusion doping would be ion implantation.

References

[1] J. E. Lilienfeld, "Method and Apparatus for Controlling Electric Currents," U. S. Patent 1745175, Jan. 1930.

[2] J. E. Lilienfeld, "Device for Controlling Electric Current," U. S. Patent 1900018, March 1933.

[3] G. C. Dacey and I. M. Ross, "Unipolar Field-Effect Transistor," Proc. IRE, **41**, p. 970-979 (Aug. 1953).

[4] D. Kahng and M. M. Atalla, "Silicon-silicon dioxide Field Induced Surface Devices," IRE-AIEE Solid-State Device Research Conference, Carnegie Institute of Technology, Pittsburgh, PA, 1960.

[5] F. M. Wanlass, "Low Stand-by Power Complementary Field Effect Circuitry," U. S. Patent 3356858, December 1967.

[6] David A. Hodges and Horace G. Jackson, *Analysis and Design of Digital Integrated Circuits, 2nd ed.*, McGraw-Hill, New York, 1988.

[7] Robert F. Pierret, *Field Effect Devices, Volume IV, 2nd ed.*, Addison-Wesley, Reading, Mass., 1990.

[8] Ben G. Streetman, *Solid State Electronic Devices, 3rd ed.*, Prentice-Hall, Englewood Cliffs, New Jersey, 1990.

[9] Else Kooi, *The Invention of LOCOS*, Institute of Electrical and Electronics Engineers, New York, 1991.

[10] R. E. Kerwin et al., "Method for Making MIS Structures," U. S. Patent 3475234, Oct. 1969.

[11] W. R. Runyan and K. E. Bean, *Semiconductor Integrated Circuit Processing Technology*, Addison-Wesley, Reading, Mass., 1990.

[12] B. E. Deal and A. S. Grove, "General Relationship for the Thermal Oxidation of Silicon," Journal of Applied Physics, **36**, p. 3770-3778 (Dec. 1965).

[13] B. E. Deal, "Standardized Terminology for Oxide Charges Associated with Thermally Oxidized Silicon," IEEE Transactions on Electron Devices, **ED-27**, p. 606-608 (March 1980).

[14] Takuo Sugano, ed., *Applications of Plasma Processes to VLSI Technology*, Wiley & Sons, New York, 1985.

[15] Werner Kern and David A. Puotinen, "Cleaning Solutions Based on Hydrogen Peroxide for use in Silicon Semiconductor Technology," RCA Review, **31**, p. 187-206 (June 1970).

[16] D. Scott Becker et al., "Effects of Various Chemistries on Silicon-Wafer Cleaning," *ACS Symposium Series 295, Microelectronics Processing: Inorganic Materials Characterization*, Lawrence A. Casper, ed., American Chemical Society, Washington, DC, 1986, p. 366-376.

[17] Don Burkman, "Optimizing the Cleaning Procedure for Silicon Wafers Prior to High Temperature Operations," Semiconductor International, **4**, p. 103-116 (July 1981).

[18] Carver Mead and Lynn Conway, *Introduction to VLSI Systems*, Addison-Wesley, Reading, Mass., 1980.

[19] R. J. Kriegler et al., "The Effect of HCl and Cl_2 on the Thermal Oxidation of Silicon," J. Electrochem. Soc., **119**, p. 388-392 (March 1972).

[20] Edmond J. Janssens and Gilbert J. Declerck, "The Use of 1.1.1.-Trichloroethane as an Optimized Additive to Improve the Silicon Thermal Oxidation Technology," J. Electrochem. Soc., **125**, p. 1696-1703 (Oct. 1978).

[21] D. W. Hess and B. E. Deal, "Kinetics of the Thermal Oxidation of Silicon in O_2/HCl Mixtures," J. Electrochem. Soc., **124**, p. 735-739 (May 1977).

[22] O. D. Trapp et al., *Semiconductor Technology Handbook*, Technology Associates, Portola Valley, CA, 1985.

[23] B. E. Deal, "The Current Understanding of Charges in the Thermally Oxidized Silicon Structure," J. Electrochem. Soc., **121**, p. 198C-205C (June 1974).

[24] Reda R. Razouk and B. E. Deal, "Dependence of Interface State Density on Silicon Thermal Oxidation Process Variables," J. Electrochem. Soc., **126**, p. 1573-1581 (Sept. 1979).

[25] Stanley Wolf and Richard N. Tauber, *Silicon Processing for the VLSI Era, Volume 1: Process Technology*, Lattice Press, Sunset Beach, CA, 1986.

[26] S. M. Sze, *VLSI Technology, 2nd ed.*, McGraw-Hill, New York, 1988.

[27] David J. Elliott, *Integrated Circuit Fabrication Technology, 2nd ed.*, McGraw-Hill, New York, 1989.

[28] T. Tran, private communication.

[29] N. G. Tarr, private communication.

[30] A. J. van Roosmalen et al., *Dry Etching for VLSI*, Plenum Press, New York, 1991.

[31] I. Fränz and W. Langheinrich, "Conversion of Silicon Nitride into Silicon Dioxide through the Influence of Oxygen," Solid-State Electronics, **14**, p. 499-505 (1971).

[32] G. Fitzpatrick, private communication.

[33] Peter Gise and Richard Blanchard, *Modern Semiconductor Fabrication Technology*, Prentice-Hall, Englewood Cliffs, NJ, 1986.

[34] J. Monkowski and J. Stach, "System Characterization of Planar Source Diffusion," Solid State Technology, **19**, p. 38-43 (Nov. 1976).

[35] Yutaka Ishikawa et al., "The Enhanced Diffusion of Arsenic and Phosphorus in Silicon by Thermal Oxidation," J. Electrochem. Soc., **129**, p. 644-648 (March 1982).

[36] T. Tran and Y. Loke, private communication.

[37] H. M. Naguib and L. H. Hobbs, "Al/Si and Al/Poly-Si Contact Resistance in Integrated Circuits," J. Electrochem. Soc., **124**, p. 573-577 (April 1977).

[38] *NanoSpec/AFT, Model No. 010-0180, Instruction Manual, Rev. 1*, Nanometrics Inc., Sunnyvale, CA, September 1982.

[39] *Alpha-Step 200 Instruction Manual*, Tencor Instruments, Mountain View, CA, April 1985.

[40] Dieter K. Schroder, *Semiconductor Material and Device Characterization*, John Wiley & Sons, New York, 1990.

[41] *Veeco FPP-5000 4-Point Probe Operation and Maintenance Manual*, Veeco Instruments Inc., Plainview, NY.

[42] J. M. David and M. G. Buehler, "A Numerical Analysis of Various Cross Sheet Resistor Test Structures," Solid-State Electronics, **20**, p. 539-543 (1977).

[43] S. T. Ahn and W. A. Tiller, "A Staining Technique for the Study of Two-Dimensional Dopant Diffusion in Silicon," J. Electrochem. Soc., **135**, p. 2370-2373 (Sept. 1988).

[44] R. Subrahmanyan et al., "Accurate Junction-Depth Measurements Using Chemical Staining," *Semiconductor Fabrication: Technology and Metrology, ASTM STP 990*, Dinesh C. Gupta, ed., American Society for Testing and Materials, 1989.

[45] C. P. Wu et al., "Techniques for Lapping and Staining Ion-Implanted Layers," J. Electrochem. Soc., **126**, p. 1982-1988 (Nov. 1979).

[46] S. J. Proctor and L. W. Linholm, "A Direct Measurement of Interfacial Contact Resistance," IEEE Electron Device Letters, **EDL-3**, p. 294-296 (Oct. 1982).

[47] John C. Irvin, "Resistivity of Bulk Silicon and of Diffused Layers in Silicon," Bell System Tech. Journal, **41**, p. 387-410 (March 1962).

[48] PDS® Phosphorus PH-900 n-Type Planar Diffusion Source Technical Data, Form A-14,081, 6/92, The Carborundum Company, Boron Nitride Division, Amherst, NY, 1992.

[49] K. Harris, "A Study of Diffusion Using Solid Source Phosphorus Pucks," EE 495 Engineering Physics Research Project, University of Alberta, Edmonton, AB, 1995.

[50] S. K. Dew, EE 619 Course Notes, University of Alberta, Edmonton, AB, 1994.

[51] H. H. Berger, "Contact Resistance and Contact Resistivity," J. Electrochem. Soc., **119**, p. 507-514 (April 1972).

[52] T. A. Schreyer and K. C. Saraswat, "A Two-Dimensional Analytical Model of the Cross-Bridge Kelvin Resistor," IEEE Electron Device Letters, **EDL-7**, p. 661-663 (Dec. 1986).

[53] Neil H. E. Weste and Kamran Eshraghian, *Principles of CMOS VLSI Design, A Systems Perspective, 2nd ed.*, Addison-Wesley, Reading, Mass., 1993.

[54] Yannis P. Tsividis, *Operation and Modeling of the MOS Transistor*, McGraw-Hill, New York, 1987.